Dissertation

Targeted quantitative proteomics in the NF-κB signalling pathway and the ubiquitin-proteasome system

zur Erklarung des akademischen Grades

(Dr. rer. nat.)

im Fach Biologie/Molekularbiologie

eingereicht an der

Lebenswissenschaftlichen Fakultät

der Humboldt-Universität zu Berlin

von

Patrick Edmund Beaudette, MSc.

Präsident der Humboldt-Universität zu Berlin

Prof. Dr.-Ing. Dr. Sabine Kunst

Dekan der Lebenswissenschaftlichen Fakultät

Prof. Dr. Bernhard Grimm

Gutachter:

1. Prof. Dr. Thomas Sommer
2. Prof. Dr. Claus Scheidereit
3. Prof. Dr. Achim Leutz

Datum der Einreichung: 26.06.2017

Datum der Promotion: 13.11.2017

Table of Contents

1 Introduction 4
1.1 The NF-κB signalling pathway 4
1.2 Canonical versus non-canonical NF-κB signalling mechanisms 5
1.3 Co-regulated signal-induced processing of both NF-κB precursors 6
1.4 The NF-κB precursors and partial proteasomal processing. 9
1.5 A brief history of mass spectrometry and protein analysis. 10
1.6 Peptide versus whole protein MS 11
1.7 Mass analyzers in proteomics. 11
1.8 Development of shotgun proteomics. 12
1.9 Stable isotopes and quantitative proteomics 13
1.10 Protein-protein interaction and post-translational modifications. 14
1.11 Undersampling and the need for targeted MS analysis. 15
1.12 Selected-reaction monitoring mass spectrometry. 17

2 Aim of this study. 18

3 Results 19
3.1 SRM-MS quantitation in the non-canonical NF-κB pathway. 19
3.1.1 Design and development of selected reaction monitoring methods: Criteria for peptide selection. 19
3.1.2 Chemical synthesis of peptide candidates permits optimization and refinement of SRM parameters. 20
3.1.3 Evaluation of synthetic peptide candidates by their MS performance 21
3.1.4 Benchmarking the performance of the SRM-MS methods 29
3.1.5 Considerations in preparing protein lysate for MS analysis. 31
3.1.6 Proof of concept: Challenging the SRM-MS methods. 33
3.1.7 Extended WT MEF stimulation. 42
3.1.8 Dynamic SILAC SRM-MS analysis of MEF LTβR stimulation time course reveals distinct precursor and product populations. 44
3.1.9 Inhibition of the ribosome, proteasome, or VCP-p97 blocks accumulation of *de novo* M-Arg-6 products. 47
3.1.10 Mathematical modelling of the LT-dependent processing of the precursors predicts signal-dependent processing of a p100-p105 complex. 50
3.1.11 Expansion of SRM-MS analysis to the broader NF-κB pathway. 56
3.2 The Protein-protein interaction network of NF-κB precursor p100. 61
3.2.1 Determination of p100 interactome by immunoprecipitation. 61
3.3 MS screening for radiation-induced interactors of NEMO/IKKγ 67
3.3.1 Optimization of immunoprecipitation and MS conditions to improve enrichment of NEMO. 68
3.4 Analyzing polyubiquitination by mass spectrometry. 73
3.4.1 Ubiquitin linkage profiling by SRM-MS. 73
3.4.2 Polyubiquitin and parkin 75
3.4.3 Ubiquitin and protein kinase A signalling. 78

4 Discussion 81
4.1 Inception and evolution of proteomics. 81
4.2 Targeted proteomics for analytically challenging scarce protein species 83
4.3 Critical facets of the SRM-MS method development process. 85
4.4 The advantages of SRM-MS over Western blotting. 87
4.4.1 SRM-MS vs. Western blotting: Quality of the assay. 88

4.4.2 SRM-MS vs. Western blotting: Quality of the results. 89
4.5 Dynamic SILAC SRM-MS analysis of LT-β R stimulated MEF cells. 91
4.6 Modelling predicts a signal-responsive precursor complex. 92
4.7 Protein-protein interaction analysis of p100 and IKKϒ/NEMO. 94
4.8 Polyubiquitination in cell signalling networks. 97
4.8.1 Linear ubiquitination and parkin. 98
4.8.2 Novel regulation of protein kinase A signalling. 98
4.9 Quantitative proteomics and unbiased biology. 99
4.10 Caveats of absolute quantitation. 100
4.11 Concluding remarks. 102

5 Summary 104

6 Zusammenfassung 105

7 Materials and methods. 107
7.1 Materials. 107
7.1.1 Chemicals. 107
7.1.2 Enzymes. 108
7.1.3 Laboratory equipment 108
7.1.4 Instrumentation. 108
7.2 Methods 109
7.2.1 Solid-phase peptide synthesis of candidates for NF-κB assay. 109
7.2.2 Quantified SpikeTides TQL peptide standards. 109
7.2.3 StageTips for peptide enrichment and desalting. 110
7.2.4 Preparation of cell pellet for enzymatic digest. 110
7.2.5 Bradford determination of protein concentration. 111
7.2.6 NanoDrop determination of peptide concentration. 111
7.2.7 SDS-PAGE preparatory gel prior to in-gel enzymatic digestion. 111
7.2.8 In-gel protein digest. 111
7.2.9 In-solution protein digest. 112
7.2.10 LC-SRM-MS analysis of NF-kB in MEF cells. 112
7.2.11 LC-MS/MS measurement of peptides derived from in-gel digestion of immunoprecipitations, and bioinformatic analysis of raw data. 113
7.2.12 In-house manufacture of reversed-phase analytical columns with emitters. 114
7.3 Collaborator contributions: cell culture, immunoprecipitations, and mathematical modelling. 114

8 Appendices 115
8.1 Abbreviations 115
8.2 List of Figures 118
8.3 List of Tables 120
8.4 Publications 121
8.5 Acknowledgements 122
8.6 Eidesstattliche Erklärung 123

9 References 124

1 Introduction

1.1 The NF-κB signalling pathway.

Investigating the NF-κB family of transcription factors is compelling due to their involvement in the regulation of a wide variety of biological activities, helping to modulate the immune response, inflammation, cell growth and tissue development according to a number of stimuli [1]. Several different signalling systems converge on the NF-κB pathway in order to elicit the desired transcriptional response, the outcome being tailored to the particular tissue with the appropriate gene regulation. For such a central system with expansive effects, it is unsurprising that faulty regulation has been linked to numerous pathologies, such as Hodgkin's lymphoma and other malignancies and disorders [2,3] [4]. A better understanding of the dynamics of NF-κB activation is therefore essential for the intelligent design of therapeutic interventions. This transcription factor family is comprised of RelA/p65, RelB, c-Rel, p50 and p52, which form various homo- and heteromeric dimers that are the functional units in gene regulation [5] [6]. Each protein shares a Rel homology domain (RHD) that mediates this dimerization, as well as binding to DNA and promoting nuclear localization. In the absence of stimulation these dimers are sequestered in the cytoplasm in an inactivated form, achieved by binding to an inhibitory protein of the IκB family, a group of proteins bearing an Ankryin Repeat Domain (ARD) that masks the nuclear localization signal [6]. The cytoplasmic IκB family includes IκBα, IκBβ and IκBε, as well as the precursor proteins p100 and p105 [7]. Regarding the precursors p100 and p105, the C-terminal half contains the inhibitory ARD, while the N-terminal half are the NF-κB transcription factors p52 and p50, respectively. Upon activation of the pathway the inhibitory IκB proteins are degraded, freeing the associated dimers to translocate to the nucleus where they can bind to promoter regions of their target genes and recruit co-activators and the transcriptional machinery. Activation of the NF-κB pathway occurs when different stimuli trigger a signalling cascade that leads to phosphorylation of the IκB molecule at specific serine residues via the IκB kinase (IKK) complex [8] [9], a critical hub comprising two types of kinases, IKKα and IKKβ, bound to a regulatory subunit IKKγ/NEMO [10,11]. Phosphorylation is closely followed by ubiquitination and proteasome-mediated degradation of the relevant IκB.

1.2 Canonical versus non-canonical NF-κB signalling mechanisms

Depending on the nature of the stimuli and the response required, two types of NF-κB pathways can be triggered, the canonical or the non-canonical, each differentiated by their mechanisms and kinetics. Canonical signalling is activated rapidly and propagates transiently until feedback mechanisms, such as NF-κB mediated upregulation of IκBs, blunt and terminate the response[12]. Canonical activation depends upon the regulatory ubiquitin binding behaviour of NEMO that activates the IKKβ kinase activity. NEMO facilitates an interaction with IκBα[13] and thereby promotes its phosphorylation, ubiquitination and ultimate degradation and the liberation of its associated dimers, typically RelA/p50.

By contrast the non-canonical signalling pathway is distinguishable by its slow and steady activation and persistence, appropriate kinetics considering the governing signals are related to development, including B-cell maturation, bone metabolism and lymphoid organogenesis[14]. Activation of the non-canonical pathway is primarily through ligand binding to the tumor necrosis factor (TNF) receptor family, such as lymphotoxin β receptor (LTβR), BAFFR, CD40, RANK and TNFR2[5,9]. Non-canonical signal propagation does not require NEMO but rather depends on the NF-κB inducing kinase (NIK), an enzyme that is suppressed at very low basal levels in the absence of stimulus by TRAF-cIAP mediated K48 polyubiquitination and proteasomal degradation[15,16]. Pathway stimulation disrupts the TRAF-cIAP complex, promoting an elevation of NIK levels by new protein synthesis[17-19]. Elevated levels of NIK result in phosphorylation of the activation loop of IKKα, which in turn phosphorylates C-terminal serine residues on the NF-κB precursor p100, triggering subsequent $SCF^{\beta\text{-}TrCP}$-catalyzed formation of K48 polyubiquitin chains[15,16]. In a very curious step, only the inhibitory ARD-containing C-terminal half of p100 is degraded by the proteasome, while the N-terminal product, now known as p52, escapes destruction and can translocate to the nucleus, usually in a dimer with RelB[20,21] . It was believed that the other NF-κB precursor, p105, undergoes spontaneous proteasomal processing to form p50, and upon canonical pathway activation, such as with lipopolysaccharide or TNFα, undergoes phosphorylation and ubiquitination events that ultimately lead to its complete degradation [22 23 24]. Signal-induced processing of endogenous p105 has not been previously observed.

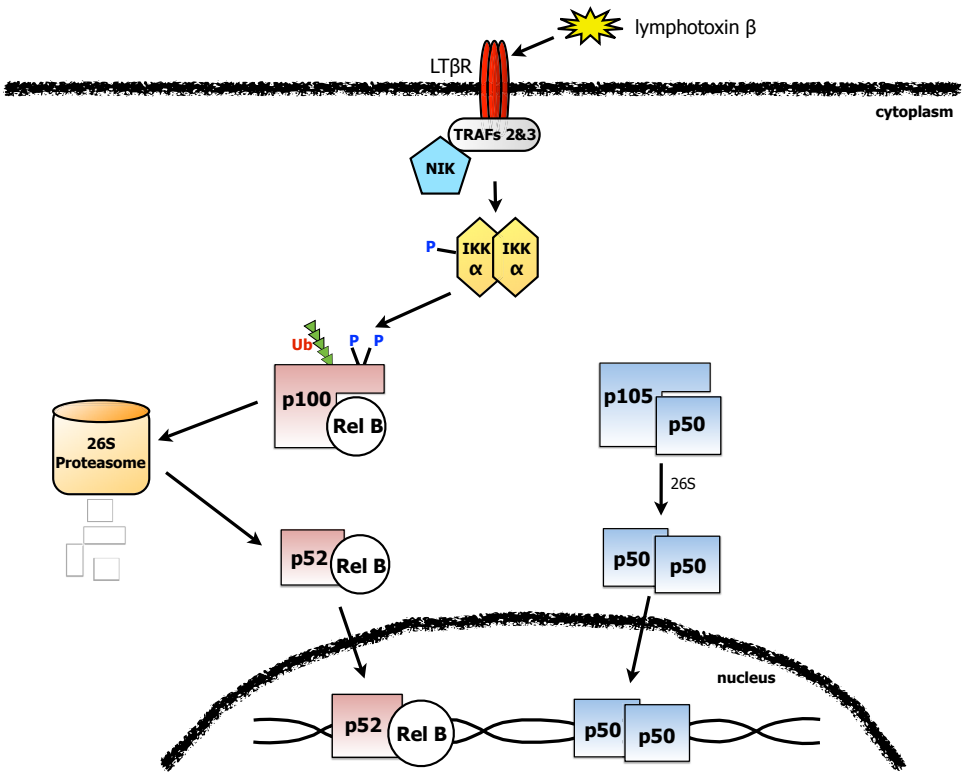


Figure 1.1 Signaling via the noncanonical NF-κB pathway.

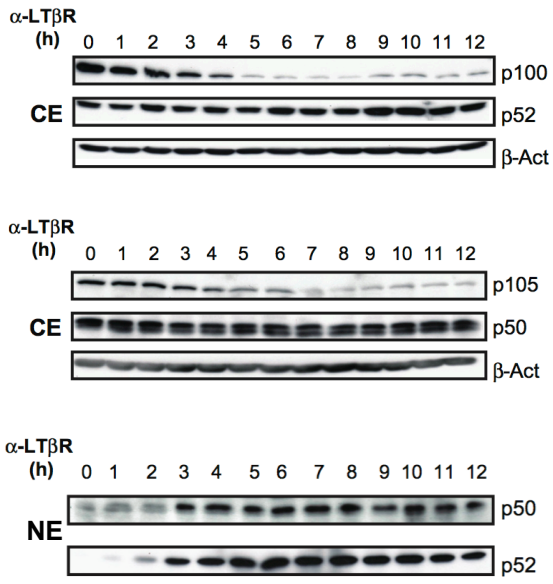
Binding of lymphotoxin β to the lymphotoxin β receptor (LTβR), leading to a cascade of NIK and IKKα-mediated phosphorylations that results in polyubiquitinated p100 being converted to the transcriptionally active p52 by a specific proteasomal truncation, with p105 being processed to p50 in parallel. Active NF-κB dimers are liberated and translocate to the nucleus to effect the transcriptional response to the external stimulus.

1.3 Co-regulated signal-induced processing of both NF-κB precursors.

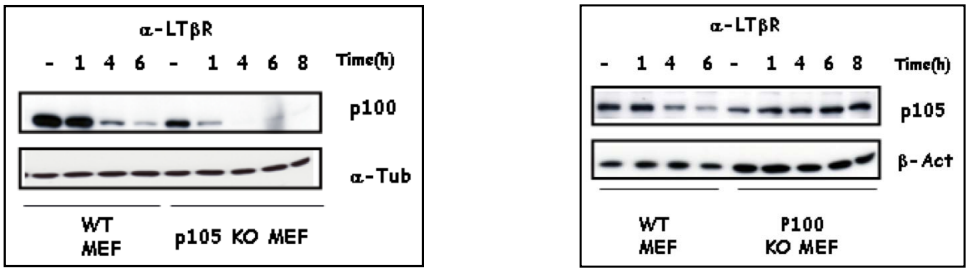
Non-canonical signalling induces the partial proteasomal processing of the precursor p100 to the product p52 via a signalling cascade relying on NIK and IKKα. We have recently published data demonstrating that the other NF-κB precursor, p105, responds in parallel to a non-canonical stimulus, an agonist of LTβR, and undergoes signal-dependent proteasomal truncation to the active product p50[25]. Figure 1.1 provides a simplified overview of non-canonical activation of the precursors p100 and p105. A western blot analysis of the precursors in cytoplasmic extract following non-canonical activation via the lymphotoxin β receptor (LTβR) is shown in Figure 1.2A, where one can observe the slow decrease of precursor band over time. Non-

canonical signal responsiveness also appears to be linked, as p105 loses its ability to respond to LTβR stimulation in murine embryonic fibroblast (MEF) cells where p100 has been knocked out, undergoing no signal-induced decrease in protein levels over time (Figure 1.2B, right panel). Conversely, when it is p105 that has been genetically deleted, the kinetics of p100 processing are enhanced, being processed to p52 more rapidly and to a greater extent (Figure 1.2B, left panel). We could show in coimmunoprecipitation studies that the two precursors routinely co-purify, suggesting a physical interaction and possible association within the context of a complex. This observation is further supported by the fact that the two species co-migrate during gel-filtration chromatography, eluting in the same high-molecular-weight fractions (Figure 1.2C). While the data in Figure 1.2 is a component of the publication stemming from my thesis project, it is introduced now as it was instrumental in shaping the scope of my work, and its inclusion here is designed to aid the reader in understanding the project's landscape upon my entry into it. A better understanding of the co-dependent processing of the NF-κB precursors may shed light on the broader regulatory aspects of non-canonical signalling.

A.



B.



C.

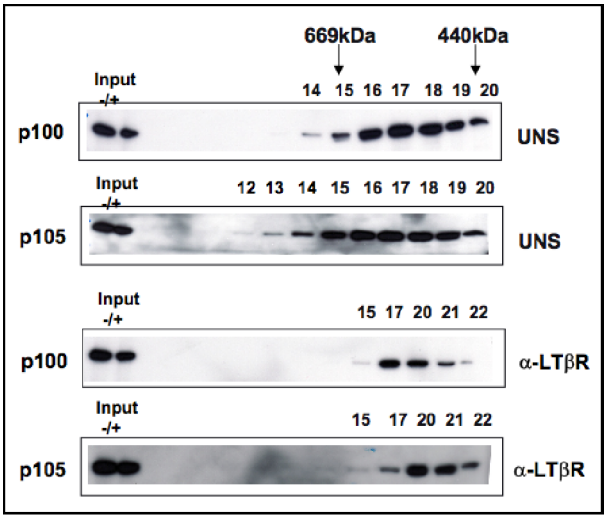


Figure 1.2 Western blot analysis of LTβR-induced MEF cells.

A. Wild-type MEF cells were induced by treatment with an anti-LTβR agonist (AC.H6) and cells sampled over 12 h. Shown are Western blots of p100/p52 (top) and p105/p50 (middle) in cytoplasmic extracts, and Western blots of p50 and p52 in nuclear extracts (bottom).

B. Wild-type and p105 KO MEF cells were induced by treatment with an anti-LTβR agonist (5G11B). Cells were sampled over 8 h and cytoplasmic extracts immunoblotted for p100 and p105.

C. Gel filtration chromatography of stimulated (α-LTβR) and unstimulated (uns) wild-type MEF lysate, with collected fractions analyzed by immunoblotting for p100 and p105.

*All MEF cell culture, LT stimulation, gel filtration and immunoblotting performed by Dr. Buket Yilmaz.

1.4 The NF-κB precursors and partial proteasomal processing.

One crucial link between the precursors is the fact that they both undergo proteasomal processing, destruction of their C-terminal half up to a specific peptide bond. The precision of this action is at odds with the usual function of the proteasome, a hydrolytic destruction chamber from which no intact protein emerges[26]. In fact, its processing of p100 and p105 is the rare instance of such a function in mammals; another transcription factor Gli3 also undergoes a proteasomal truncation. Its analogue in Drosophila is a transcription factor of the Cubitus interruptus (Ci) family undergoing a proteasome-mediated C-terminal truncation that serves to reverse its role from being a transcriptional activator to a repressor[27]. In a process with obvious parallels to the NF-κB precursor, Ci155 is cleaved to form Ci75 after undergoing phosphorylation and ubiquitination. Regulation of this proteasome-mediated action has been linked to Ter94, an ATPase associated with diverse cellular functions. (AAA ATPase). The mammalian ortholog of Ter94 is called VCP/p97, a hexameric ATPase capable of generating mechanical force from ATP hydrolysis and acting on its substrates through a family of UBX-domain containing adapter proteins[28,29]. It plays a role in different degradative pathways, such as ubiquitin-fusion degradation (UFD) and endoplasmic-reticulum-associated degradation (ERAD) [30], and is also essential for reformation of the nuclear envelope during the cell division cycle[31]. Within the NF-κB universe itself, p97 has been linked to the degradation of IkBα in the canonical pathway[32], a process involving sequential phosphorylation and ubiquitination events. Investigating p97's effect on the precursor processing in the non-canonical NF-κB signalling pathway is therefore a worthy venture, and the recent discovery of a reversible, competitive small molecule inhibitor of p97 known as DBeQ provides a convenient tool for better exploring any mechanistic links[33].

Up to this point the study of partial proteasomal processing of the NF-κB precursors has been achieved largely through traditional molecular biology approaches, with stimulus-derived concentration changes detected by Western blotting methods. The present project proposes to adapt targeted proteomic methods based on selected-reaction monitoring (SRM) mass spectrometry (MS) in order to measure temporal concentration changes of the NF-κB precursors and their products in the presence of non-canonical pathway stimulation, namely activation of the LTβR. SRM-MS is used in conjunction with nanoflow high-performance liquid chromatography (HPLC) that enables absolute quantitation of peptides in a complex biological matrix[34]. Copy numbers of proteins can be determined and precise concentration fold-changes

calculated. SRM-MS offers a very high degree of specificity and sensitivity, a vastly improved mode of detection that has only been made possible in the last decade due to advances in the field of mass spectrometry-based proteomics[35].

1.5 A brief history of mass spectrometry and protein analysis.

When J.J. Thomson discovered electrons with his cathode ray tube and measured their mass-to-charge ratio 120 years ago[36], the era of mass spectrometry was born. In the beginning its use was relegated largely to physicists and was essential in several fundamental scientific findings, such as the discovery of stable isotopes in the 1920s[37]. Up until the 1950s the major practical application was found in the petroleum industry to profile hydrocarbons in distillation fractions[38], with the first peptide sequencing paper appearing in 1959 to usher in a biochemical age for this technology[39]. In this paper the non-volatility of peptides was finally overcome through a reductive derivatization of carbonyl groups into polyaminoalcohols, which could then form gas-phase ions after harsh, fragment-inducing electron impact (EI) ionization. While this ionization was successful only with di- and tripeptides it was still a major breakthrough. When cleverly combined with Edman degradation and early DNA sequencing, whole protein sequences were elucidated in this way, albeit with painstaking effort and time [40-42]. Further advances in ionization technology, from chemical ionization (CI) [43] to fast atom bombardment (FAB) [44] enabled larger peptides to be ionized without derivatization. From this early point, the protease trypsin was used to convert full-length proteins into lower molecular weight, MS-amenable peptides. The development of electrospray ionization (ESI) at Yale University in 1989 by John Fenn and colleagues was a massive leap forward and the most important milestone yet in protein mass spectrometry, a method to gently convert large biomolecules into gas-phase ions directly from the liquid phase, with minimal fragmentation [45,46]. ESI represented significant gains in ionization efficiency and ion transfer into the mass analyzer and, importantly, ionizing directly from the liquid phase meant that ESI could be coupled directly to an HPLC separation. The molecular weight limitations imposed by other ionization methods were also effectively removed, as the ESI process generates multiply-charged ions, meaning even large biomolecules have a mass-to-charge ratio compatible with the detection limits of most mass analyzers.

1.6 Peptide versus whole protein MS.

While the ESI process can ionize an intact protein[47], digesting whole proteins into peptides offers superior analytical benefits. Peptides are much more uniform in their physicochemical profiles, particularly their size, solubility and hydrophobicity[48]. The majority of a protein's peptides will not be modified by a glycosylation chain, a common variable-length post-translational modification (PTM) resulting in a single species presenting many mass-to-charge ratios, a process dividing the intensity across many channels and reducing sensitivity. In terms of MS analysis, peptides have superior ionization potential, leading to improved signal response and sensitivity compared to proteins, and their fragment spectra are easier to interpret[49]. The complexity of a whole cell tryptic digest is staggering, a solution with potentially a million unique analytes. Fortunately peptides are also very amenable to chromatography, with reverse-phase C18 columns routinely employed to resolve them in-line with the MS analysis[50].

1.7 Mass analyzers in proteomics.

As ionization technology was advancing, so too were the mass analyzing technologies at the heart of a mass spectrometer. Room-sized magnetic sector instruments[51] were superseded by triple-quadrupole instruments that placed a collision cell between two mass analyzing quadrupoles, facilitating mass analysis on both intact and fragmented molecules[52,53]. The first quadrupole selects an ion of interest and filters out all remaining signals. The transmitted ion enters the collision cell where it strikes molecules of nitrogen or argon at high speed, causing internal molecular bonds to break. Ion fragments then enter the third quadrupole where a mass spectrum is recorded. Called tandem mass spectrometry, or MS/MS, it provides spectra rich in structural information. In the case of peptides, fragmentation along the peptide bond backbone means that, with some experience, the amino acid sequence of the peptide can be inferred. Other mass analyzers used in biochemistry include time-of-flight (TOF), converting an ion's transit time in a drift tube to a mass-to-charge ratio, offering high resolution and fast scanning across the mass range[54]. Ion trapping mass analyzers could enrich ions both before and after fragmentation and thereby increased the sensitivity of the scans, and multiple levels of fragmentation (MS^n) allow for deeper structural characterization of molecules[55]. The current apex of ion-trapping technology is the Orbitrap, an analyzer first commercially

released in 2000 that now dominates the proteomic landscape[56,57]. It consists of a spindle-shaped electrode within a cylindrical electrode. Trapped ions orbit around the spindle with a motion dependent on their mass-to-charge ratio. This motion is detected as an image current which is then converted by Fourier transform into a mass spectrum[58]. The quadrupole and the Orbitrap are the two most important mass analyzers being used in proteomics. Table 1.1 below compares their relative merits[48].

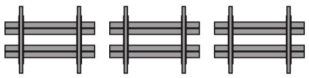
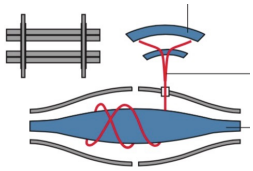
Property	Quadrupole 	Orbitrap 
m/z range	100-1250	50-1800
Resolution	500-1500	15,000-250,000
Mass accuracy	20 ppm	0.5 ppm
Speed	Slow, 2000 Da/s	Fast, 12,000 Da/s
Ion transmission efficiency	High	High
Expense	$	$$$

Table 1.1 Comparison of triple-quadrupole and Orbitrap mass analyzers.

1.8 Development of shotgun proteomics.

The completion of the Human Genome Project ushered in an era where the completed genomes of humans and other organisms made accessible all of the protein sequences encoded in their DNA, whether or not these gene products had been previously observed [59,60]. Using the genome, one can make a database of the complement of peptides all of its proteins would yield when subjected to a tryptic digest. Experimental MS spectra of peptides derived from an experimental tryptic digest of an unknown protein could be matched to those predicted by the DNA sequence in a method called Peptide Mass Fingerprinting (PMF), the concept being that a tryptic digest of a given protein yields a mix of tryptic peptide masses uniquely

characteristic of that protein[61-63]. While far faster than the classical Edman degradation, this technique worked only for relatively pure samples, such as a band excised from a polyacrylamide gel. A technique applicable to complex samples of many digested proteins in solution that relied on MS/MS spectra of tryptic peptides soon followed. Peptide Fragment Fingerprinting (PFF) matches the experimentally observed fragments of a tryptic peptide in an MS/MS spectrum with a database of *in silico* predicted MS/MS spectra derived from a database[64,65]. This approach, termed "shotgun" or "bottom-up" proteomics could be applied to complex samples such as whole cell lysates digested with trypsin. Shotgun proteomics completely transformed protein biology and led to an explosion of discovery, as identifying proteins no longer required tedious efforts to first isolate them for traditional sequencing.

1.9 Stable isotopes and quantitative proteomics.

Early shotgun proteomics efforts were satisfied with identifying as many proteins in a sample as possible, but soon strategies to reliably quantify proteins arising from different cells or samples dominated the literature[66]. Understanding how protein expression levels are altered between tissues, according to an external stimulus or in a disease state, holds great promise for illuminating mechanistic questions. However it is vital that such differences were not simply an artefact of the analysis, such as sample preparation or instrument response, and incorrectly assigned a biological meaning. An internal standard is needed that can normalize a peptide's intensity to account for technical variations that often arise. Tagging or labelling molecules with stable isotopes proved to be an excellent answer, adding neutrons to peptide molecules so a portion of the hydrogen (^{1}H), carbon (^{12}C), nitrogen (^{14}N) or oxygen (^{16}O) atoms are replaced by a heavier stable isotope (^{2}H or D, ^{13}C, ^{15}N, and ^{18}O) [67-70]. Assuming sufficient resolution, two peptides with identical amino acid sequences but slightly different neutron content could be resolved in a mass spectrum, giving two mass-to-charge ratios. Aside from this incremental mass difference, the two species are otherwise identical, co-eluting in a chromatographic separation and ionizing with the same potential. As an internal standard, isotope-coded peptides can correct for technical variations in an LC-MS/MS analysis - injection volume, instrument response, chromatography – and highlight true differences in protein levels between samples founded in biology. Incorporation of stable-isotopes can be done both chemically and metabolically. Chemical tags typically target a peptide's primary amine groups,

found at the N-terminus and on lysine side-chains, on either intact proteins or on tryptic peptides. This can be done inexpensively for two or three parallel samples by dimethylating with formaldehyde via a reductive amination process[71,72], or using commercial products offering higher levels of multiplexing with isobaric tags that release quantitation marker fragment ions upon MS/MS fragmentation[73,74]. For cells grown in culture, heavy-labelled amino acids can be incorporated into all proteins using the cell's own translation machinery in a process called Stable Isotope Labeling of Amino acids in Cell culture (SILAC) [75-78]. SILAC exploits the principle that lysine is an essential amino acid that the cell obtains from the growth media. While arginine is not an essential amino acid, biosynthesis is limited and it is mostly obtained from diet [79]. Substituting heavy lysine and arginine amino acids into the culture media, SILAC ensures that every tryptic peptide will be labelled at its C-terminus. By labelling early in the processing workflow, technical variations related to chemical tagging can be avoided. The most recent method in quantitation gaining traction in the field does not use stable isotopes at all. Label-free quantitation (LFQ) relies on multiple repeat measurements of samples and is valid only when the technical inter-run variations are statistically less than any biologically-derived differences between samples being compared[80,81]. For each sample in the set, peptide signals in the form of high-resolution extracted-ion chromatograms are integrated and averaged with the other peptides of a protein that have been detected across all samples, allowing the calculation of normalization factors for each sample to correct for technical variations in the analysis. The appeal of LFQ is the absence of additional chemical manipulations of the sample, and its application for biological systems where SILAC cannot be employed, for example in animal models and clinical samples[82].

1.10 Protein-protein interaction and post-translational modifications.

With a sound strategy for normalization, quantitative shotgun proteomics is a very powerful approach for not only identifying thousands of proteins present within a given proteome, but for gaining biological insights into how proteins are differentially expressed under certain conditions such as disease. Coupled to immunoprecipitations, quantitative proteomics can be used to identify those proteins specifically enriched in a sample through direct or indirect association to the bait protein[83-85]. This provides evidence of an interaction that may or may not have

functional implications for both proteins, and can help discover protein complexes and define their membership [86] [87]. Similar strategies can be used to examine differences in levels of a post-translational modification. Ciechanover wrote that "Timely and selective protein modification is a crucial feature of many modes of protein regulation" [88]. This is particularly true with phosphorylation and ubiquitination, and there exist many examples in the literature identifying and quantifying these modifications [89] [90] [91] [92] [93] [94]. Polyubiquitination, the extension of the primary ubiquitination site by multiple ubiquitin molecules, plays an important role in signalling pathways, including both canonical and non-canonical NF-κB activation [95] [96]. The mode by which one ubiquitin molecule is attached to the next in the chain determines the fate of the substrate on which the chain is formed. Eight possibilities exist for the second and subsequent ubiquitin molecules to connect to the one before it, via primary amine groups on the seven internal lysines and the N-terminus of ubiquitin. Each linkage has a unique topology and particular linkages have come to be associated with particular functions, the classical being K48-linked polyubiquitin [97] [98]. This modification signals the destruction of a target protein and leads to its recruitment and hydrolysis at the proteasome [99]. Knowledge of the character of a polyubiquitin chain is important to understanding the purpose of a ubiquitination event on a specific target protein as well as exploring how the global linkage landscape is altered by perturbing the system [100] [101]. Linear polyubiquitin (or M1) chains form between the N-terminal amine of a conjugated ubiquitin and the C-terminal acid group of the incoming ubiquitin molecule [101]. This chain type has been shown to play an essential role in certain modes of NF-κB activation [102] [103], and depends on the 'linear ubiquitin assembly complex' (LUBAC) a ligase complex comprised of Sharpin, HOIL-1L and HOIP.

1.11 Undersampling and the need for targeted MS analysis.

Continuous innovation and development in mass spectrometry results in new instrumentation that collect spectra at a higher rate and achieve greater sensitivity being released every couple of years. Despite these advances, the sheer complexity of proteomic samples remains a significant analytical challenge, one that is further compounded by the vastly different expression levels of proteins across the proteome [104]. Termed the 'dynamic range problem', it refers to the fact that some proteins are expressed at high levels across tissues (e.g. mammalian VCP/p97 is

thought to comprise 1% of cytosolic protein [105]) while others are present at vanishingly low amounts [106] [107]. In mammalian cells, this range spans 6 orders of magnitude: for every copy of a rare protein, there are a million copies of the most abundant protein. In human plasma the problem is even more extreme, with a span of 10 orders of magnitude, or 10 billion of the most common protein molecules for every one of the least common [108]. Like the proverbial needle lost in a haystack, the peptides from the most abundant proteins dominate the sample while the least abundant are easily lost to obscurity [109]. In 2010, 50% of all MS/MS spectra found in public data repositories represented only 82 human proteins [110]. Unlike genomic research, there is no amplification scheme available to boost levels of rare proteins until they are in the detection range. As a consequence, these rare proteins are underrepresented in datasets, overlooked by instrumentation working beyond its capacity. The most common type of shotgun proteomics measurement is called a data-dependent analysis (DDA) and is structured as a Top 10 experiment [48] [111]. An initial full spectrum MS scan identifies the ten most intense peptide peaks currently eluting from the column at a given time. Following this survey scan, ten MS/MS scans of these selected peptides are acquired before the cycle repeats itself. To avoid redundancy a given signal cannot be selected again for MS/MS for a short duration to allow the elution to finish. One issue with this approach is that ionization suppression effects between co-eluting peptides means that many weaker signals are never selected for MS/MS, with a bias towards the most abundant proteins that have higher MS intensities. Another is that despite the advances in instrumentation, even at a lofty MS/MS acquisition rate of 200 scans per minute, it is mathematically impossible to sequence every peptide from complex whole-cell digest samples. This problem is called undersampling and comes with the drawback that shotgun data sets suffer from a reproducibility problem: two back-to-back measurements of the same sample will not cover all of the same proteins or peptides [112]. Naturally it is the least abundant proteins where this effect is felt most acutely, in which category the transcription factors such as the NF-κB family tend to fall. This poses a challenge for quantitative proteomics experiments when the same proteins must be measured across a large number of samples, such as those derived from a clinical study or from a time-course experiment.

1.12 Selected-reaction monitoring mass spectrometry.

A different approach to MS analysis is required to alleviate this issue. When the analytical goal is to reproducibly quantify a pre-selected group of proteins in a sample and not the entire proteome, it is appropriate to employ a Targeted Proteomics approach, namely SRM-MS [35] [113]. Unlike a data-dependent shotgun analysis, targeted proteomics focuses the instrument's attention on a predetermined group of peptides with known MS/MS fragmentation patterns that represent the proteins of interest in the study. The LC separation component does not change, but the type of MS mass analyzer does. LC-SRM-MS utilizes the mass filtering properties of a triple-quadrupole (QQQ) MS instrument [114] to selectively monitor a specific peptide analyte. The first quadrupole filters all masses except that corresponding to the target peptide in its known charge state. The selected fragment ion is fragmented in the second stage by high-energy impacts with a collision gas [115]. Finally the third quadrupole filters and measures specific fragment ions derived from that peptide. This precursor ion/fragment ion pair is called a transition and several transitions per peptide are measured to improve the method's selectivity. A series of transitions corresponding to multiple peptides are cycled through repeatedly and record multiple data points as a peptide elutes from the chromatographic column. The resulting chromatographic peaks can be integrated and the area under the curve used for relative quantitation. When heavy-isotope-labelled internal standard peptides are spiked into the sample, their signal forms the basis for absolute quantitation of the peptide analysis. The present project aims to adapt targeted proteomic methods based on SRM-MS detection in order to measure the concentrations of the NF-κB precursors and their processing products following LTβR stimulation.

2 Aim of this study.

Activation of the non-canonical NF-κB signalling pathway is known to result in proteolytic processing of p100 by the proteasome to form p52. Our publication [25] has expanded this mechanism to include a stimulus-dependent processing of p105 to its transcriptionally active form p50, a process tightly-coupled to p100. The aim of the current study is to examine the processing and degradation of endogenous NF-κB precursors p100 and p105 in a time-resolved manner in murine embryonic fibroblasts (MEFs). To this end, traditional immunology-based detection methods are replaced with a targeted proteomics approach employing selected-reaction monitoring mass spectrometry (SRM-MS) to facilitate absolute quantitation of both the precursors p100 and p105 and their respective products p52 and p50. The strategy relies on the incorporation of stable isotopes into synthetic peptides for quantitation, and into proteins through metabolic labeling of cells in culture to distinguish pre- and post-stimulation populations of proteins. Ultimately this absolute, time-resolved quantitative data will be used to challenge various mathematical models of this pathway activation mechanism to support or refute the hypothesis that the NF-κB precursors p100 and p105 form a complex that responds to the lymphotoxin β receptor activation signal in a concerted manner. Building upon this work, the protein-protein interaction network of p100 will be investigated to identify other proteins with a functional role in the activation mechanism, particularly with respect to the proteasomal processing step.

The canonical NF-kB pathway, in particular the DNA damage response, will be studied with non-targeted (i.e. data-dependent) proteomic methods. The protein-protein interaction network of endogenous IKKY/NEMO will be unravelled in an effort to identify cytoplasmic partners of this important signalling hub that bind to an IKKY/NEMO complex in response to irradiation.

Additionally, both targeted and non-targeted proteomics methods will be applied to determine the nature of polyubiquitin linkages present in a specific system. Here the goal is attaining a better understanding of parkin's role in anti-apoptotic activity as well as mechanisms governing the regulation of protein kinase A (PKA) signalling

3 Results.

3.1 SRM-MS quantitation in the non-canonical NF-κB pathway.

3.1.1 Design and development of selected reaction monitoring methods: Criteria for peptide selection.

Selection of an appropriate peptide to act as proxy for the quantitation of its parent protein is a critical first step in the development of a selected reaction monitoring method. Foremost, the chosen peptide(s) must represent an amino acid sequence that is unique, termed proteotypic. A proteotypic peptide is necessary for the unambiguous measurement of the target protein. Additional criteria impacting peptide selection are MS sensitivity, chromatography and number of potential isoforms. The ideal peptide exists in one form, meaning a single charge state and no chemical or post-translational modifications. Multiple isoforms of a peptide complicate quantitation, as multiple internal standards are required to tally the contributions made by each. Moreover as each form is likely to possess a discrete mass, the presence of multiple signals essentially divides the ion current amongst all of them, lowering the overall intensity and reducing sensitivity. The list below summarizes criteria most important for peptide selection [116]:

Peptide properties to avoid:

- Methionine residues.
- Annotated or suspected sites of post-translational modification.
- Sequence motifs triggering missed cleavage events e.g -KR, -KK.
- Short hydrophilic peptides.
- Long hydrophobic peptides.

Oxidation of methionine, post-translational modification and missed cleavages all result in multiple isoforms of the same peptide. Short, hydrophilic peptides are undesired as they are not well retained by reverse-phase chromatography and can elute too early in the gradient. Conversely, long hydrophobic peptides may elute too

late, and solubility problems may cause difficulty when using synthetic peptides as an internal standard.

3.1.2 Chemical synthesis of peptide candidates permits optimization and refinement of SRM parameters.

Selection of peptides according to the above guidelines is helpful for filtering out potentially problematic sequences. *In silico* prediction tools [117,118] can be helpful in narrowing the field of options, but empirical evidence is far more valuable as the MS performance of a particular peptide is challenging to predict.

Solid-phase synthesis of peptides is achieved through coupling amino acid derivatives in the required order on a resin support using Fmoc protecting groups to block reactive moieties in the side-chains so that amino acids are joined together through the backbone carboxylic acid and primary amine groups. The schematic in Figure 2.1 shows an Fmoc-protected amino acid derivative being added to a growing peptide chain bound to a solid support.

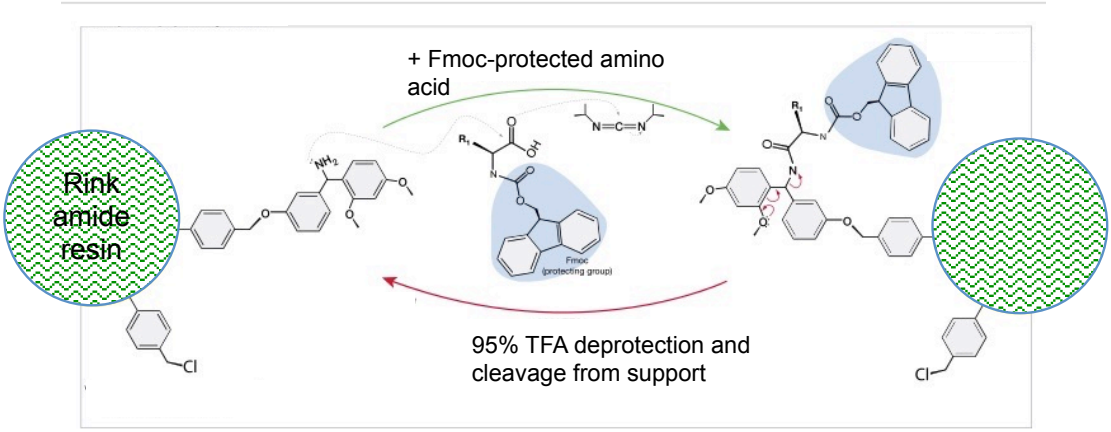


Figure 2.1 Solid-phase peptide synthesis using a Rink amide support. (adapted from Dan Cojocari/Wikipedia)

For our purposes a Rink amide (hydroxymethyl-polystyrene) resin was employed, the resultant peptide products having an amidated C-terminus rather than a free acid group [119]. This trade-off is weighed against the gentler cleavage conditions (95% trifluoroacetic acid, TFA) to efficiently release the peptides from the resin. While the amidated synthetic peptide is 1 Da heavier than its native counterpart it behaves

virtually identically. Importantly, the peptide will adopt the same charge state and fragmentation patterns and therefore provide an opportunity for tuning and optimizing the MS instrument parameters in order to maximize the performance of the SRM method.

Following the solid-phase synthesis procedure, the resins were washed thoroughly with Ethanol before elution with the TFA cleavage buffer. The peptides were precipitated from the eluate with cold methyl-t-butyl ether to afford several milligrams of white to off-white solid. A purity of 90% was assumed and the peptide products dissolved in 30% acetonitrile containing 0.1% formic acid to make stock solutions of 1 mM. The reaction products were evaluated in the next step by MS characterization on a triple-quadrupole mass spectrometer, and the candidates with the most intense MS signal response were selected for further development.

3.1.3 Evaluation of synthetic peptide candidates by their MS performance.

The motion of a positively-charged gas-phase peptide ion within the vacuum of a mass spectrometer is manipulated by application of negative electrical potential and radiofrequency waves. Analogous to how light is bent and focused in a microscope with lenses and mirrors, the MS uses ring electrodes and quadrupoles to selectively transmit and focus beams of gas-phase ions. The schematic in Figure 2.2 illustrates the ion path of a triple-quadrupole mass spectrometer.

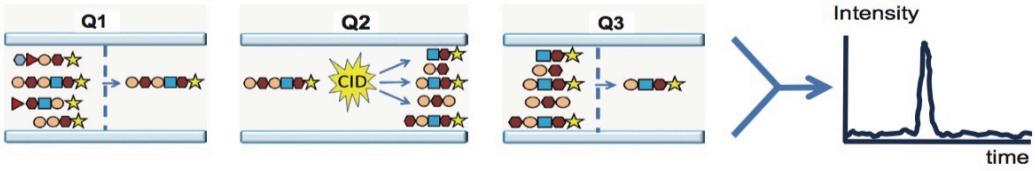


Figure 2.2 Schematic representation of triple-quadrupole (QQQ) mass spectrometer.

Mass filtering at Q1 selects the precursor peptide ion, which is transmitted to Q2 where it is fragmented by high-energy collision with nitrogen gas molecules. Mass filtering of the selected fragment mass in Q3 completes the circuit, generating a measureable ion current that is plotted.

The triple quadrupole instruments used for this work, Sciex 5500 or 6500 Q-Traps, have a third quadrupole that can also function as a linear ion trap, trapping and

enriching precursor or fragment ions until they are scanned out to the detector. This functionality was used in method development to amplify the signal and increase the signal-to-noise. All peptides derived from the solid-phase synthesis were subjected to the following MS characterization. A synthetic tryptic peptide from p100 (TPSPSGSLLR) will be used as an example to illustrate the process of method development. The first step is confirmation of the expected mass of the synthetic peptide. The formula: m/z = (M+ nH)/n gives the mass (m) to charge (z) ratio for a given charge state, where n is the number of additional protons. Tryptic peptides generally have a minimum of two basic (i.e. proton accepting) sites, the N-terminal amine and the basic side-chain of the C-terminal lysine or arginine residue. In electrospray MS at acidic pH both sites are charged, meaning the majority of tryptic peptides have a charge state of 2. Figure 2.3 shows an MS1 spectrum for the peptide TPSPSGSLLR collected by scanning a 100-1000 mass range at 1000 Da/s in the linear ion trap; the view has been zoomed to 30 Da of the x-axis to better show detail. The peptide has a calculated molecular weight of 1014.55 Da, so the observed value of 507.4 Th indicates that this peptide has two additional protons resulting in the double charge. (1013.55+2)/2 = 507.8 (observed mass offset from calculated mass due to deviation in calibration of the linear ion trap). The resolution of this scan is insufficient to fully resolve the isotopes of the peptide.

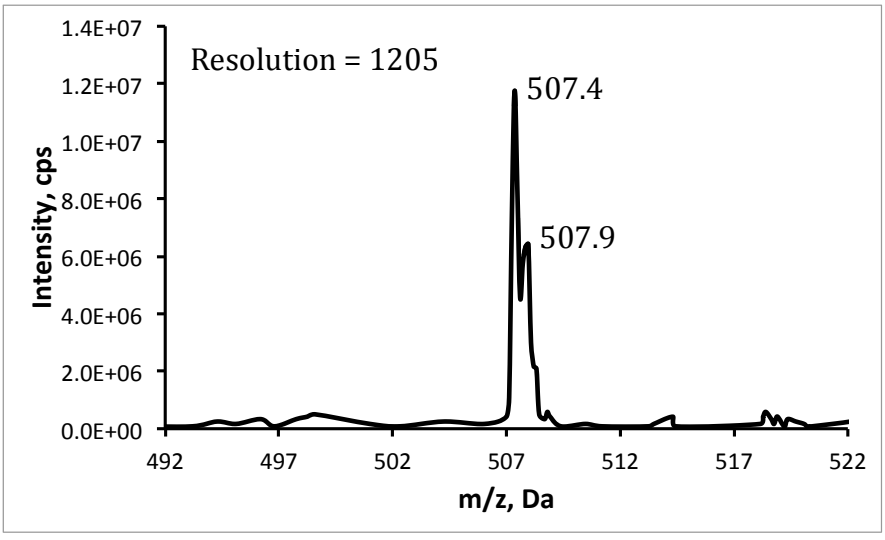


Figure 2.3 Enhanced MS spectrum of the p100 peptide TPSPSGSLLR.

In Figure 2.4 an enhanced resolution scan resolves the isotopes of the peptide to the baseline, with the higher-mass peaks corresponding to ^{12}C atoms replaced by one or

two ^{13}C atoms, the proportion dictated by the natural abundance of this isotope, 1.1%. “Enhanced resolution” is achieved by collecting ions in the linear ion trap and scanning a narrow mass range at a slow speed. The spacing of the isotope peaks provides another way to determine the charge state of the molecule by counting the number of isotopes occurring within a single unit of the m/z scale.

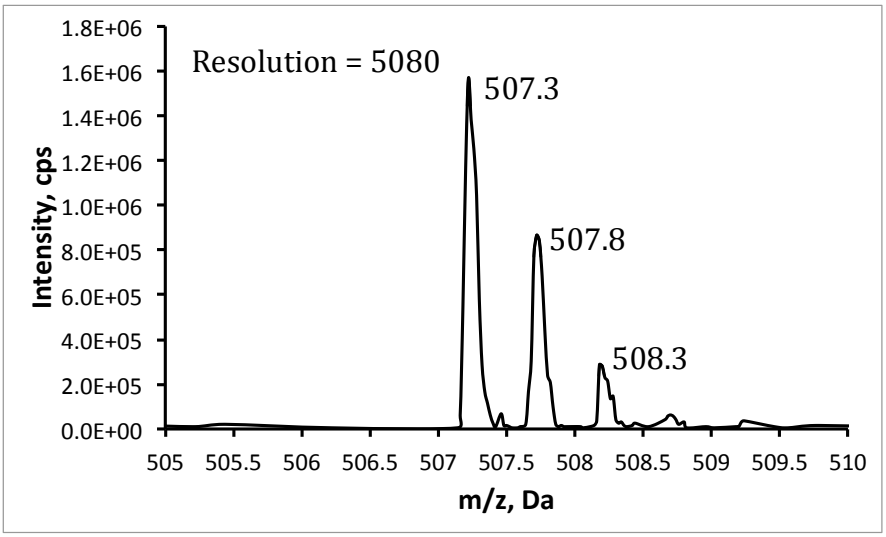


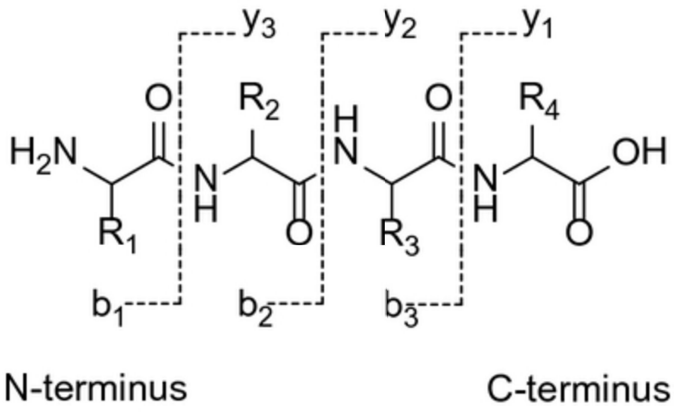
Figure 2.4 Enhanced Resolution spectrum of the p100 peptide TPSPSGSLLR.

The MS1 measurements complete the first entry in the SRM method, the observed precursor mass to be isolated by quadrupole 1. An MS/MS scan is performed next to establish the collision-induced dissociation (CID) fragmentation pattern of the peptide. Figure 2.5 shows an MS/MS spectrum of the precursor with an m/z of 507.4, the location of which is indicated on the spectra with an arrow. Collision with an N_2 molecule splits the peptide into two pieces and results in singly-charged fragments flying through to the detector. When possible, incorporating fragments into the SRM method with m/z values greater than the precursor is advantageous, as it eliminates the possibility of singly-charged background ions interfering with the measurement. This is illustrated more dramatically in Figure 2.10, an MS/MS spectrum of p105 peptide VGADLSLLDR. One can clearly observe that below the precursor m/z of 529.8 the spectra is populated by more fragment ions of a higher intensity than in the higher mass ranges.

With the successful acquisition of a good quality MS/MS spectrum, the fragment ions to be filtered in quadrupole 3 can be selected. From Figure 2.5A, the fragments

corresponding to y5, y6, y7 and y8 are at higher m/z than the precursor and have a high intensity, making them ideal candidates to use in an SRM transition.

A.



B.

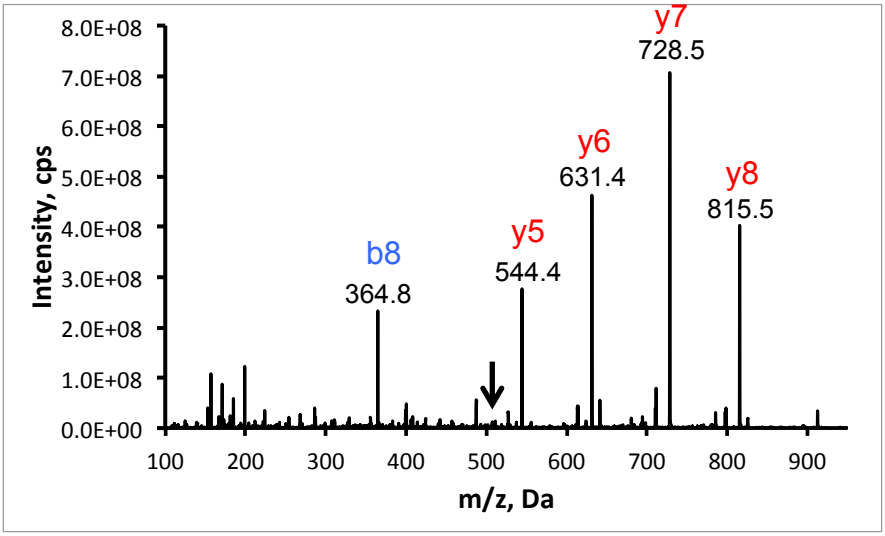


Figure 2.5

A. Schematic of a generic peptide with naming convention of N-and C-terminal fragment types following breakage of the indicated peptide bond.

B. Product ion (MS/MS) spectrum of of p100 peptide TPSPSGSLLR with major fragments labeled according to convention. Arrow indicates m/z of precursor ion.

While a precursor and fragment mass are the minimum requirement of an SRM method, the existence of synthetic standards permits a more thorough method development by optimizing two key instrument parameters, the declustering potential (DP) and the collision energy (CE). Tuning these can allow for the maximum ion current through the system and therefore the greatest sensitivity, essential for the detection of low-abundance species. The declustering potential is a voltage applied at the orifice of the machine, where the electrosprayed ions at atmospheric pressure enter into the vacuum of the instrument, and it functions to break up the clusters of ions. Minimizing clusters is akin to maximizing the transmission of single ions, meaning the maximum amount of precursor ion can be transmitted through the first quadrupole. A declustering potential ramp from 50 to 300V was applied to the SRM transitions and the ion transmission monitored. Figure 2.6 shows the individual curves collected for each transition across this range, with the optimal value defined as the highest intensity, otherwise interpreted as the maximum ion current. One observes that the curves are relatively flat but that the maxima offer a 20-30% increase in signal strength relative to the mimima. For the p100 peptide TPSPSGSLLR a value of 150 V was selected as a common DP value for all transitions, indicated by a vertical dashed line in Figure 2.6.

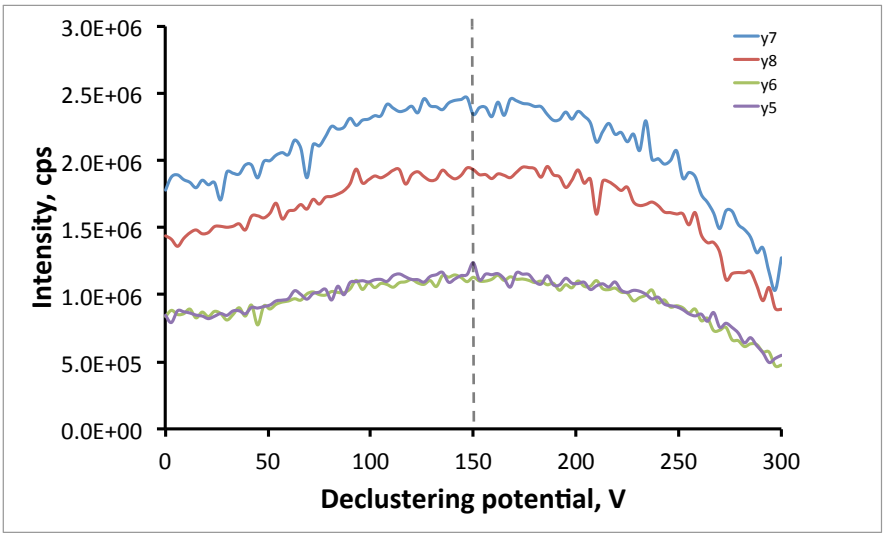


Figure 2.6 Declustering potential optimization.

A solution of the p100 peptide TPSPSGSLLR was infused at 5 µL/min into the Sciex 5500 Q-trap triple-quadrupole MS while the voltage applied at the orifice is ramped between 0 and 300 V in 5V increments during continuous acquisition of the indicated SRM transitions. The vertical dashed line is the DP value selected for all measurements.

Tuning of the collision energy offers the greatest opportunity to increase the method sensitivity. Below the optimal value a portion of the precursor molecule remains unfragmented and therefore cannot contribute to the fragment current detected beyond Q3. Above the optimal value the higher energy imparted to the precursor can lead to more than one fragmentation event per molecule. The intensity of the monitored fragment ions could lose intensity due to such a secondary fragmentation, reducing the ion current and sensitivity. The collision energy was ramped from 5 to 50 eV and the intensity for each transition measured across the range, with Figure 2.7 showing just how crucial tuning this parameter is; only a 5 eV shift to either side of the maxima corresponds to loss of intensity of approximately 25%. The dashed gray line in the figure represents the collision energy calculated using a linear equation as in Skyline [120]. While it is a reasonable estimation of the optimal value, it is clear that it oversimplifies fragmentation behaviour by only taking into account the precursor's mass and charge state.

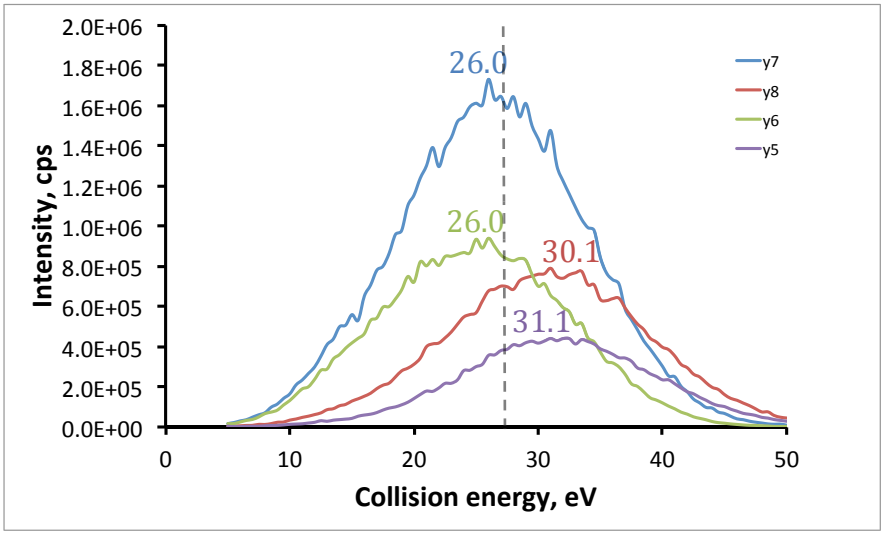


Figure 2.7 Collision energy optimization for p100 peptide TPSPSGSLLR.

A solution of the p100 peptide TPSPSGSLLR was infused at 5 µL/min into the Sciex 5500 Q-trap triple-quadrupole MS while the kinetic energy imparted to the precursor ion is ramped between 5 and 80 eV in 1 eV increments during continuous acquisition of the indicated SRM transitions. The dashed vertical line is the calculated collision energy value.

The MS characterization of the peptides from p105, p52 and p50 were carried out in the same manner, with further examples given below for the p105 peptide VGADLSLLDR (Figure 2.8 – 2.12). Table 2.1 indicates the sequence of each peptide, its position within the sequence, the precursor and fragment masses included in the method as well as the optimized collision energies.

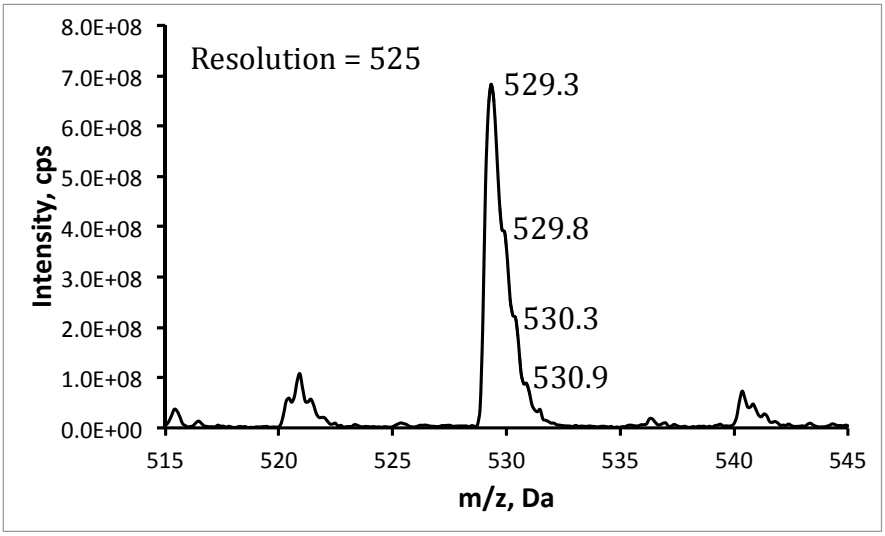


Figure 2.8 Enhanced MS scan of p105 peptide VGADLSLLDR.

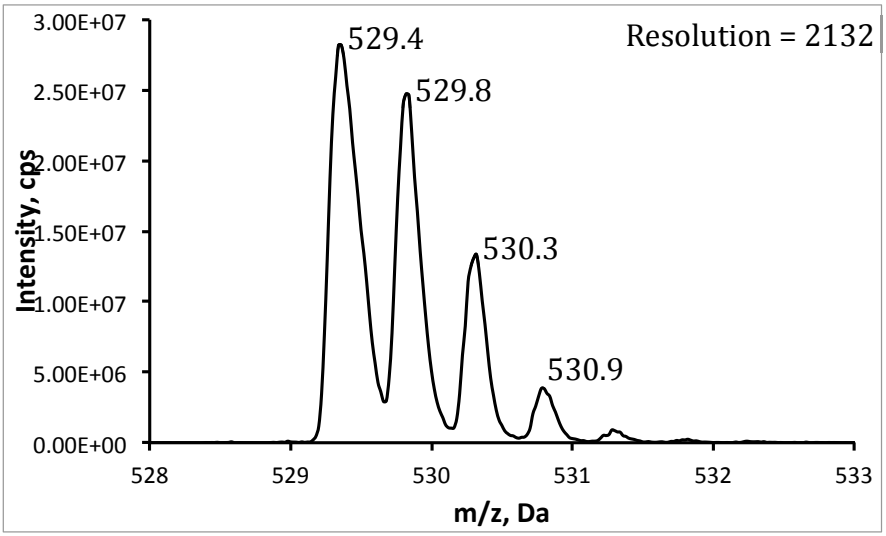


Figure 2.9 Enhanced resolution scan of p105 peptide VGADLSLLDR.

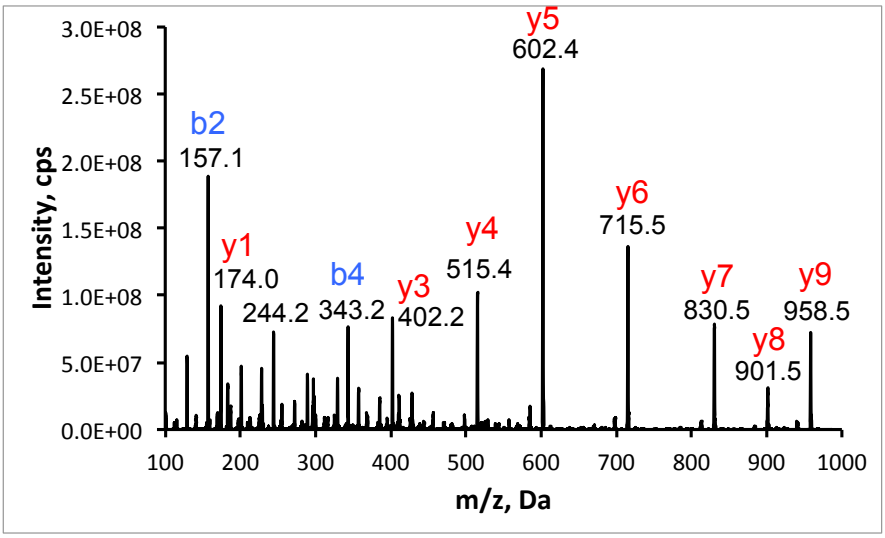


Figure 2.10 Product ion (MS/MS) spectra of p105 peptide VGADLSLLDR.

Product ion (MS/MS) spectrum of p100 peptide VGADLSLLDR with major fragments labeled according to convention. Note increase in singly-charged background ions below precursor m/z of 529.4.

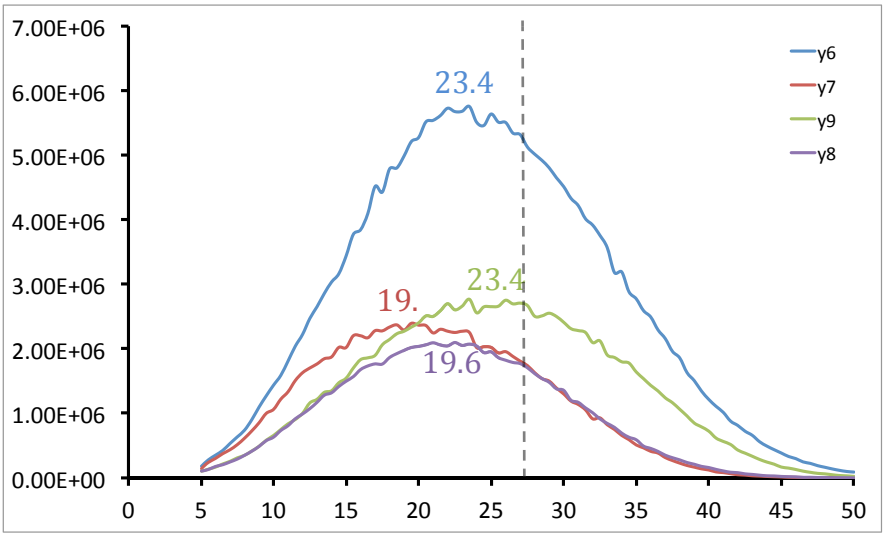


Figure 2.11 Collision energy optimization for p105 peptide VGADLSLLDR.

A solution of the p105 peptide VGADSLLDR was infused at 5 µL/min into the Sciex 5500 Q-trap triple-quadrupole MS while the kinetic energy imparted to the precursor ion is ramped between 5 and 80 eV in 1 eV increments during continuous acquisition of the indicated SRM transitions.

Figure 2.12 Schematic of tryptic peptides and their relative sequence position chosen for SRM-MS analysis of NF-kB precursors and products.

Peptide	Protein	Sequence position	SRM transitions; Fragment ID, (CE in eV)
QYAIVFR	p52	284-290	448.75 > 534.33; y4 (20) 448.75 > 421.25; y3 (21) 448.75 > 605.37; y5 (26)
TPSPSGSLLR	p100	810-819	507.78 > 729.43; y7 (26) 507.78 > 816.46; y8 (25) 507.78 > 632.37; y6 (31)
VFETLEAR	p50	147-154	482.75 > 718.36; y6 (21) 482.75 > 589.32; y5 (23) 482.75 > 865.43; y7 (20)
VGADLSLLDR	p105	600-609	529.79 > 603.34; y5 (23) 529.79 > 959.51; y9 (20) 529.79 > 831.45; y7 (22)

Table 2.1 SRM transitions for MS analysis of p100/p52 and p105/p50 protein species.

3.1.4 Benchmarking the performance of the SRM-MS methods

The next phase of method development involves measuring both the native and isotopically-heavy (Arg 10) peptide in a complex background of a whole-cell tryptic digest. The SRM-MS analysis of peptides is always coupled to a chromatography step. Most typically, a solution of tryptic peptides is injected onto a reversed-phase column and a gradient of increasing acetonitrile elutes peptides largely according to their hydrophobic properties. The SRM signal is recorded over time and manifests itself as a peak, with the integrated area under the curve a function of the amount of peptide injected. The ideal internal standard is a peptide analogue with an identical amino acid sequence but with additional neutrons incorporated in the lysine and

arginine residues that are by definition at the C-terminus of a tryptic peptide. Such a molecule is physicochemically indistinguishable from the native peptide, eluting at the same time and fragmenting in the same way, but resolved into a distinct signal by the MS. When a fixed amount of IS is injected together with the samples, the IS peak area can be utilized to normalize the signal of the native peptide and facilitate relative quantitation between samples. When the quantity of IS is known precisely, the signal can be used to apply an absolute quantitation of the native species. Such normalization is important in LC-MS experiments due to the number of variables that can waver between injections: volume of sample injected by the autosampler, drift in instrument response, and ionization suppression.

For the SRM-MS method to be a viable quantitation tool, an experiment was designed to demonstrate that the MS detector responds linearly to increasing amounts of peptide. High-purity synthetic peptides were purchased from JPT Peptide Technologies GmbH (Berlin, Germany) serially-diluted in a fixed background of a tryptic digest of MEF whole cell lysate. Such a digest presents a worst-case scenario with respect to sample complexity; an entire proteome cleaved into hundreds of thousands of unique peptide species. At any given point during the gradient elution, hundreds of peptides may be co-eluting, creating an environment of competition in the electrospray ionization process. A higher affinity for protons for some peptide suppresses the ionization of co-eluting species, preventing their conversion to gas-phase ions and reducing sensitivity for that particular target. Figure 2.13 plots the ratio of the two signals, dividing the area under the curve (AUC) of the isotopically-heavy (with a $^{13}C_6{}^{15}N_4$ composition, or H-Arg-10) internal standard by the AUC of the native ($^{12}C_6{}^{14}N_4$ composition, or L-Arg-0) peptide. Note that this is a reversal of the usual arrangement, where a fixed amount of heavy peptide is used to quantify the variable amounts of native peptide present in biological samples. The plots demonstrate several important details: the SRM method succeeds in detecting each of the targets in the complex sample, even down to 1 fmol of peptide injected in a complex background. Linear regression analysis of the plots gives R^2 values approaching 1, indicating that the instrument responds linearly over this range of concentrations. Finally, these plots are helpful in determining an appropriate amount of internal standard to include in future measurements of MEF cells. The linearity of p105's MS signal response ($R^2 = 0.94$) is not as convincing as for the other peptides ($R^2 \approx 0.99$). This can be offset by carefully choosing the amount of internal standard to spike in. For an ideal quantitation the amount of internal standard is approximately

equal to the amount of native peptide present in the sample, resulting in a peak area ratio of 1. When both analyte and standard belong to the same order of magnitude, the linearity of signal response is less strained and the internal standard can be used to effect a single-point calibration. For all of the peptide targets, an internal standard amount of 5 fmols was chosen for the quantitation of MEF-derived samples as this satisfied the ratio criteria and a single concentration simplifies the sample preparation workflow. The SRM methods designed for the quantitation of NF-κB family members p100/p52 and p105/p50 have met the analytical requirements set forth and are ready to be tested in the context of larger biological experiments.

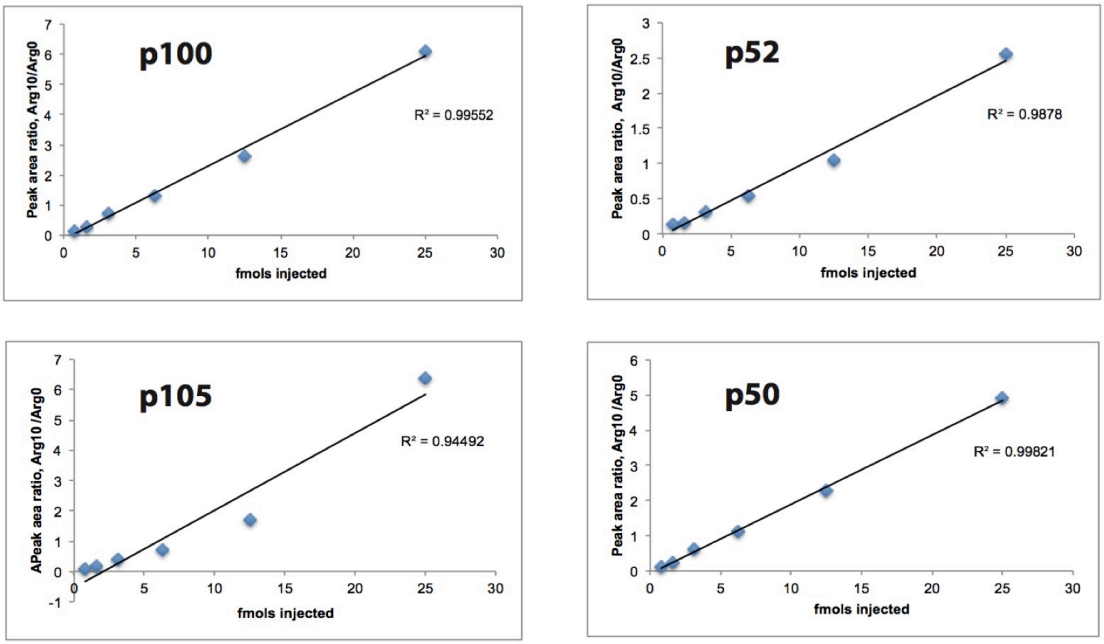


Figure 2.13 Signal response curves for NF-κB peptides.

Calibration curve prepared by spiking heavy (H-Arg-10) SpikeTide reference peptides into a fixed background of a tryptic MEF digest. Injected samples were resolved on a nanoflow chromatographic system (Eksigent) and on-line SRM-MS analysis was performed on a Sciex 5500 Q-Trap triple-quadrupole in positive ion mode Plotted are the ratio of the H-Arg-10 peptide signal divided by the endogenous L-Arg-0 peptide signal.

3.1.5 Considerations in preparing protein lysate for MS analysis.

Sample preparation in proteomics is a key step that will impact the success of any downstream measurement. For samples derived from cell culture, the objective is to

produce a reasonably concentrated protein solution free of undesirable cellular components. These include:

1. Genomic DNA: This must be mechanically or enzymatically sheared into shorter (<500) base pair oligonucleotides or the solution becomes viscous. This viscosity reduces the fidelity of any liquid transfer steps, and lower diffusion rates decrease the efficiency of any chemical or enzymatic treatments. Additionally, intact DNA is highly-charged and can retain some peptides and proteins through ionic interactions.
2. Lipids and small-molecule metabolites: When these compounds are not the target of analysis they only serve to create an additional analytical burden by contributing to sample complexity and ion suppression, and can negatively impact chromatographic behaviour (see Figure 2.23, p50 measurement following Wessel-Fluegge treatment)
3. Detergents: While included in many molecular biology protocols to aid in the lysis and dissolution of plasma membranes, their presence during the electrospray process lowers surface tension and prevents formation of droplets. Detergents that are polymeric in nature (e.g NP-40, Triton X-100) can be a pervasive contaminant, dominating spectra and obscuring weaker peptide signals via ionization suppression effects across much of a gradient elution.

In addition to excluding unwanted components, it is necessary to prevent post-lysis alterations to the proteome, a common problem when the compartmentalization of the cell is disrupted, leading to non-specific proteolysis or other post-translational modifications. Inclusion of phosphatase and deubiquitinase inhibitors (NaV, NaF, NEM) are tolerated though any resulting chemical modifications they cause must be taken into consideration for the mass shift they induce in a target. Protease inhibitors must be avoided, as they will severely inhibit the tryptic digest. A complete denaturation of the proteome can achieve the same goal, preventing random background proteolysis by disrupting the structure of the proteases.

In the simplest form of proteomic sample preparation, in-solution digest, cells are lysed in a buffered solution of 6M urea/2M thiourea with the aid of vortexing and ultrasonication, the latter serving to shatter the genomic DNA as well. There is minimal loss of protein material, with the exception of membrane-bound proteins that are notoriously difficult to solubilize and are typically underrepresented in such a

preparation. Cell lysate can also be filtered using a gel matrix, such as in sodium dodecyl sulfate-polyacrylamide electrophoresis (SDS-PAGE). Small molecules (<500 Da) are readily removed and the gel-embedded proteins can be stably manipulated by chemical modification and digestion. Tryptic digestion of the reduced and denatured proteome typically proceeds with fewer missed cleavages than an in-solution preparation, but extracting the peptides for MS analysis following the digestion is problematic, with estimates of losses ranging up to 90% [121].

The Wessel-Fluegge method [122] is a liquid-liquid extraction protocol designed to partition cellular biomolecules between polar (methanol/water) and non-polar (chloroform) phases. Protein and DNA precipitates align along the interface between the two immiscible liquid phases, while small-molecule metabolites and lipids partition to the polar and non-polar phases, respectively. Though some losses occur due to incomplete precipitation and re-solubilization, the Wessel-Fluegge technique combines the high-yield of the in-solution digest method with the cleaner background of the gel preparation. Changing to Wessel-Fluegge from the in-solution digest protocol was important in remedying an issue with the measurement of p50. When prepared by in-solution digest, the measured peak was weak, broad and tailing (Figure 2.23), a challenging peak to quantify accurately as the peak boundaries are less sharp and the wide elution window is more likely to include interfering co-eluters. The drastic improvement in peak shape and intensity seen in Figure 2.21 after a Wessel-Fluegge preparation is clear. One can speculate about the deleterious role that small molecules and lipids may have played in altering the chromatographic behaviour, but what is important is that a narrower peak coupled to a higher intensity makes the assay for p105 detection both more selective and sensitive.

3.1.6 Proof of concept: Challenging the SRM-MS methods.

The dynamics of LT-βR-triggered non-canonical NF-κB signalling have previously been explored by using Western blotting techniques. Figure 1.2 showed the effect of stimulation on the levels of p100 over time, as measured in cytoplasmic extract. After 4 h of LTβR stimulation there is a sharp reduction in the intensity of the p100 band, even more pronounced at the 6h timepoint. A novel finding preceding this work is that p105 also responds to the non-canonical stimulus, the amount reducing markedly on a similar timescale as with p100. In the cytoplasmic compartment, the levels of p52 and p50 do not appear to change. When p105 is absent from the cell due to a

knock-out, p100 levels appear to decline at a faster rate and to a lower endpoint, with the band disappearing after 4h of stimulation.

In this proof-of-concept experiment, the questions asked of the SRM-MS methodology are as follows:

1. How selective is SRM-MS i.e. are there any interferences or artefacts influencing the measurement?
2. Are the same trends observed for p100/p52 and p105/p50 that are seen in the Western blot analysis?
3. Is SRM-MS detection as sensitive as Western blotting?

As mentioned, a tryptic digest of a whole cell protein extract has a staggering complexity. The possibility exists of a co-eluting peptide presenting a similar mass in Q1 and yielding fragments of similar mass. The mass filtering is relatively coarse, isolating a 0.7 m/z window for transmission to Q2, so the selectivity achieved with such a complex sample is imperfect. Co-eluting interferences might be dealt with by changing the chromatographic condition, manipulating the gradient's duration and profile to achieve a separation of the two species. If the interference cannot be completely separated and forms a shoulder with the target peak, splitting the peak at the valley when integrating is sufficient for quantitation.

MEF cell lines in which the various NF-κB precursors' and products' are not formed due to the applied knock-out provide an opportunity to examine the selectivity of the SRM-MS measurement; the absence of the target will make any co-eluting interference easier to observe. Please note that all cell material was provided by my collaborator Dr. Buket Yilmaz as pellets. The $p100^{-/-}$ and $p105^{-/-}$ MEF cell lines have the C-terminal coding region truncated but leave an intact N-terminal sequence of the precursor genes, so p52 and p50 are present. $Nfkb2^{-/-}$ and $Nfkb1^{-/-}$ are complete knockouts where neither the precursor nor the product are present in the cell.

The $Nfkb2^{-/-}$ MEF cell line was probed by the p100 and p52 SRM methods. A heavy (H-Arg-10) internal standard was included to act as a positive control and elution time marker, pinpointing the precise elution time of the L-Arg-0 peptide. In Figure 2.14, the H-Arg-10 peptide is observed to elute at 19.1 min, as indicated by the three monitored transitions co-eluting as characteristic nesting peaks at that point.

Examining the same retention time window in the L Arg-0 trace of a wild-type MEF (Figure 2.15), the nesting peaks are observed at the same timepoint, while in the p100 KO MEF (Figure 2.16) there is no corresponding signal at this retention time, indicating both the expected absence of this species in the sample and demonstrating the selectivity of the SRM-MS detection method. All of the traces shown below show that for each SRM transition being monitored, there is a significant ion current present during much of the gradient, with distinct chromatographic peaks indicating an eluting peptide that shares a precursor and fragment mass in common with the target peptide. This underlines the value for using a heavy internal standard to confirm the elution time.

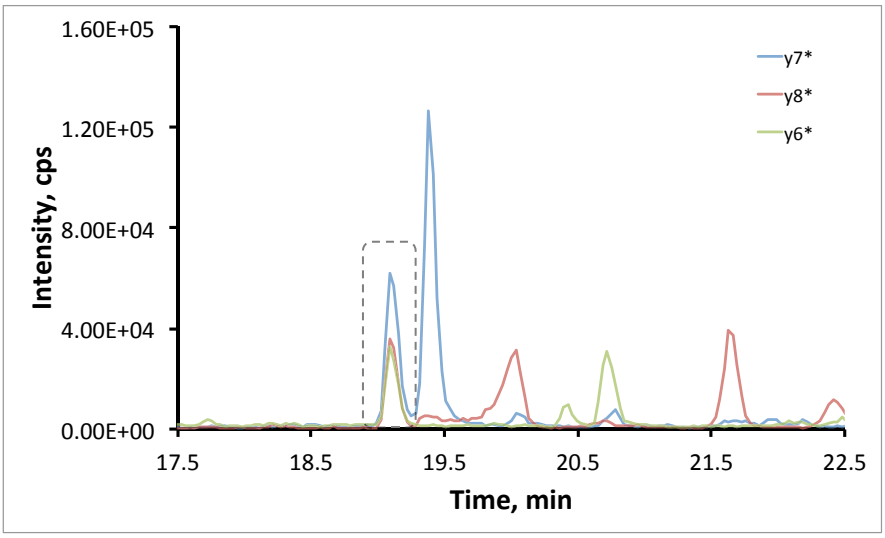


Figure 2.14 SRM-MS measurement of H-Arg-10 p100 peptide in wild-type MEF.

Five fmols of H-Arg-10 SpikeTide internal standard in a fixed background of 5 µg of a tryptic digest of wild-type whole-cell murine embryonic fibroblasts measured by SRM-MS on a 5500 Q-Trap. The dashed box indicates the SRM signals for H-Arg-10 TPSPSGSLLR.

*MEF cell pellets provided by Dr. Buket Yilmaz.

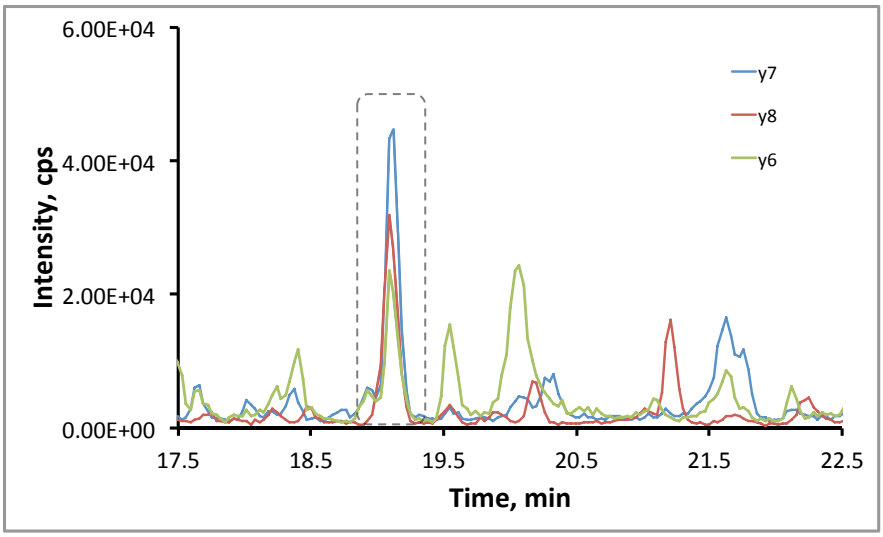


Figure 2.15 SRM-MS measurement of L-Arg-0 p100 peptide in wild-type MEF.

Five µg of a tryptic digest of wild-type whole-cell murine embryonic fibroblasts and measured by SRM-MS on a 5500 Q-Trap. The dashed box indicates the SRM signals for the L-Arg-0 peptide TPSPSGSLLR.

*MEF cell pellets provided by Dr. Buket Yilmaz.

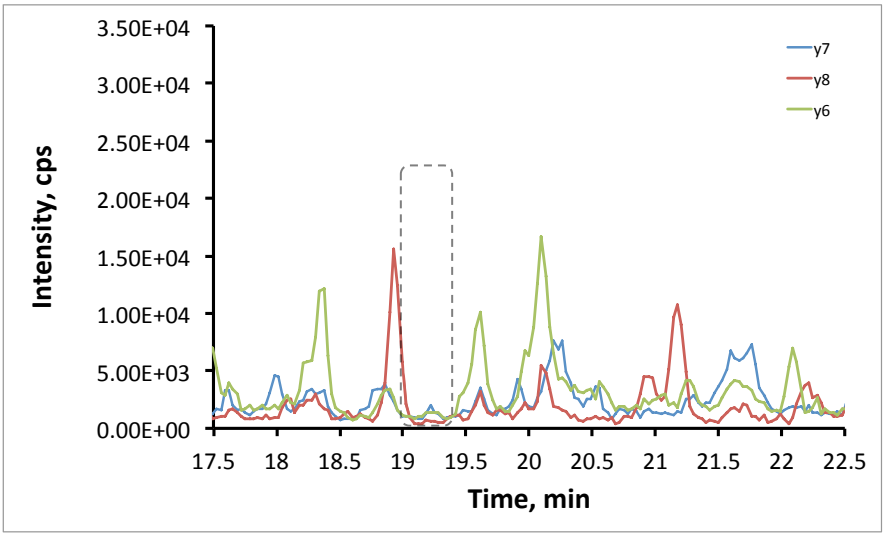


Figure 2.16 SRM-MS measurement of L-Arg-0 p100 peptide in $p100^{-/-}$ MEF.

Five µg of a tryptic digest of $p100^{-/-}$ whole-cell murine embryonic fibroblasts measured by SRM-MS on a 5500 Q-Trap. The dashed box indicates the established elution window for the L-Arg-0 peptide TPSPSGSLLR.

*MEF cell pellets provided by Dr. Buket Yilmaz.

Similarly for the other targets, the knock-out cell lines facilitated a demonstration of the selectivity of the SRM-MS methods. Figures 2.17 through 2.23 illustrate this in the detection of the peptides QYAIVFR (p52), VGADLSLLDR (p105) or VFETLEAR (p50). The absence of signal during the expected elution windows in the knockout cell lines is confirmation that the measurement of p100 and p52 is not subject to background interference co-eluting at the same time. The methods p105 and p50 were also proven selective in this manner; in both cases no signal was found in the retention window indicated by the elution of the spiked-in heavy peptide standard.

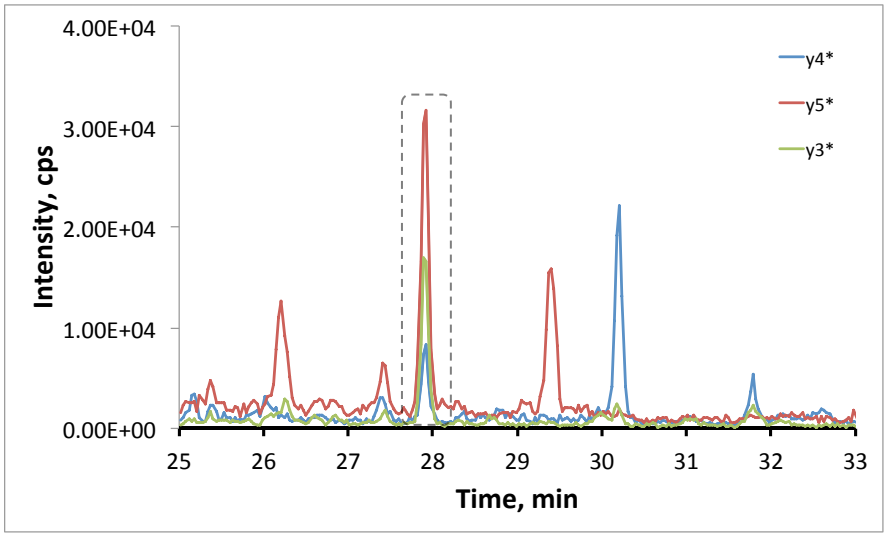


Figure 2.17 SRM-MS measurement H-Arg-10 p52 peptide in Nfkb2$^{-/-}$ MEF.

Five fmols of H-Arg-10 SpikeTide internal standard in a fixed background of 5 µg of a tryptic digest of nfkb2$^{-/-}$ whole-cell murine embryonic fibroblasts measured by SRM-MS on a 5500 Q-Trap. The dashed box indicates the SRM signal for H-Arg-10 QYAIVFR.

*MEF cell pellets provided by Dr. Buket Yilmaz.

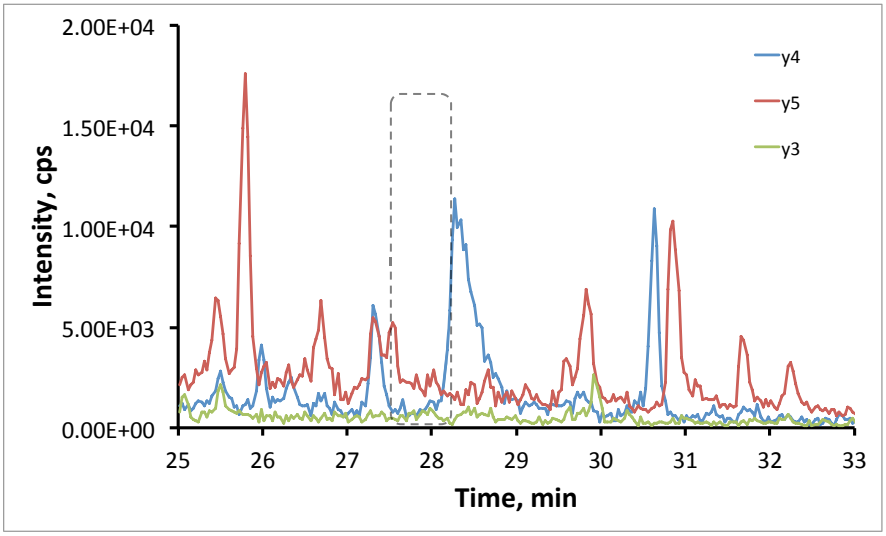


Figure 2.18 SRM-MS measurement of L-Arg-0 p52 peptide in Nfkb2$^{-/-}$ MEF.

Five µg of a tryptic digest of nfkb2$^{-/-}$ whole-cell murine embryonic fibroblasts measured by SRM-MS on a 5500 Q-Trap. The dashed box indicates the established elution window for the L-Arg-0 peptide QYAIVFR.

*MEF cell pellets provided by Dr. Buket Yilmaz.

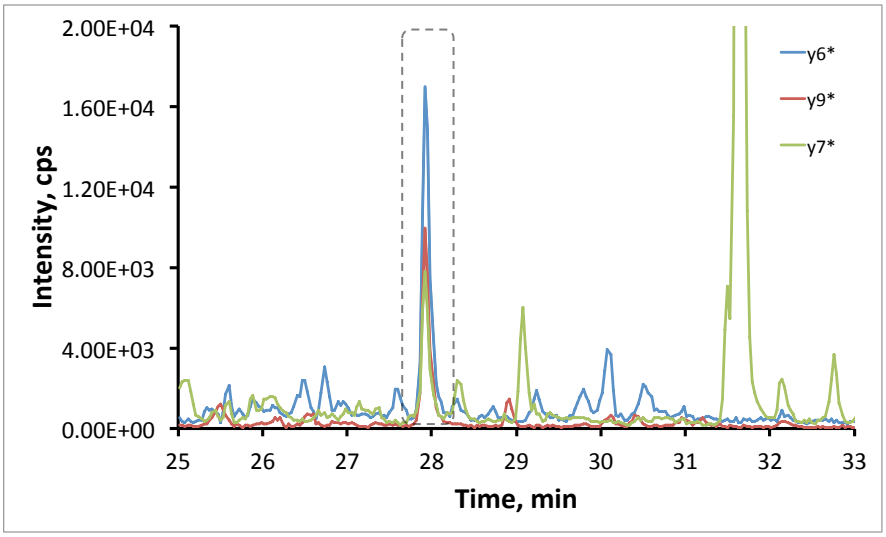


Figure 2.19 SRM-MS measurement of H-Arg-10 p105 peptide in wild-type MEF.

Five fmols of H-Arg-10 SpikeTide internal standard in a fixed background of 5 µg of a tryptic digest of wild-type whole-cell murine embryonic fibroblasts measured by SRM-MS on a 5500 Q-Trap. The dashed box indicates the SRM signals for H-Arg-10 VGADLSLLDR.

*MEF cell pellets provided by Dr. Buket Yilmaz.

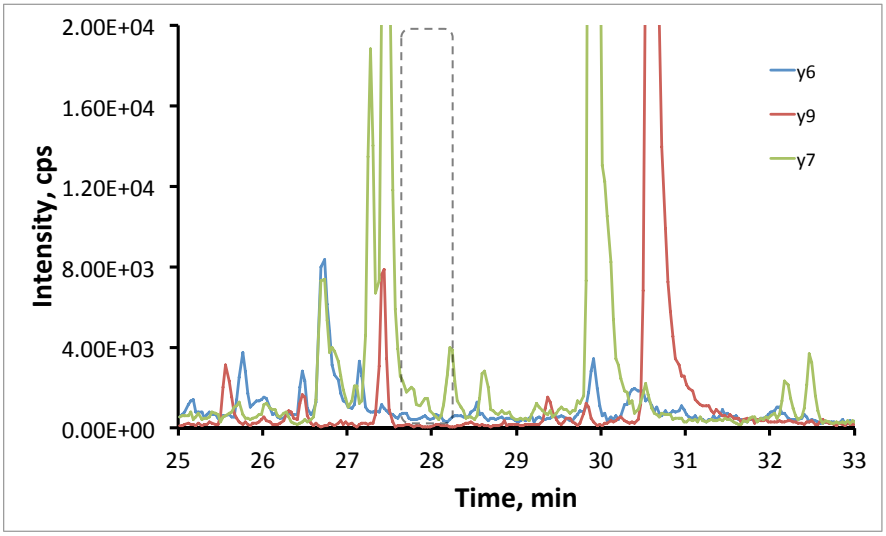


Figure 2.20 SRM-MS measurement of L-Arg-0 p105 peptide in p105$^{-/-}$ MEF.

Five µg of a tryptic digest of p105$^{-/-}$ whole-cell murine embryonic fibroblasts measured by SRM-MS on a 5500 Q-Trap. The dashed box indicates the established elution window for the L-Arg-0 peptide VGADLSLLDR.

*MEF cell pellets provided by Dr. Buket Yilmaz.

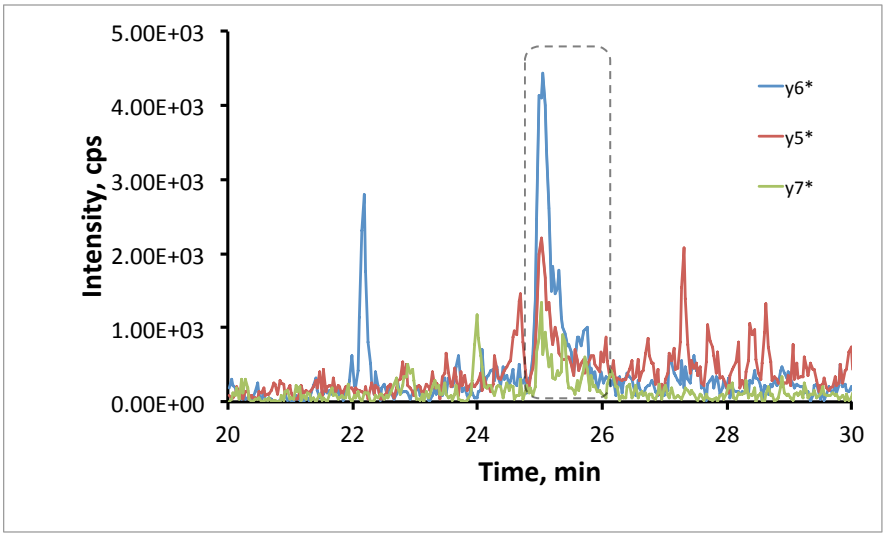


Figure 2.21 SRM-MS measurement of H-Arg-10 p50 peptide in Nfkb1$^{-/-}$ MEF.

Five fmols of HArg-10 SpikeTide internal standard in a fixed background of 5 µg of a tryptic digest of wild-type whole-cell murine embryonic fibroblasts measured by SRM-MS on a 5500 Q-Trap. The dashed box indicates the SRM signals for H-Arg-10 VFETLEAR.

*MEF cell pellets provided by Dr. Buket Yilmaz.

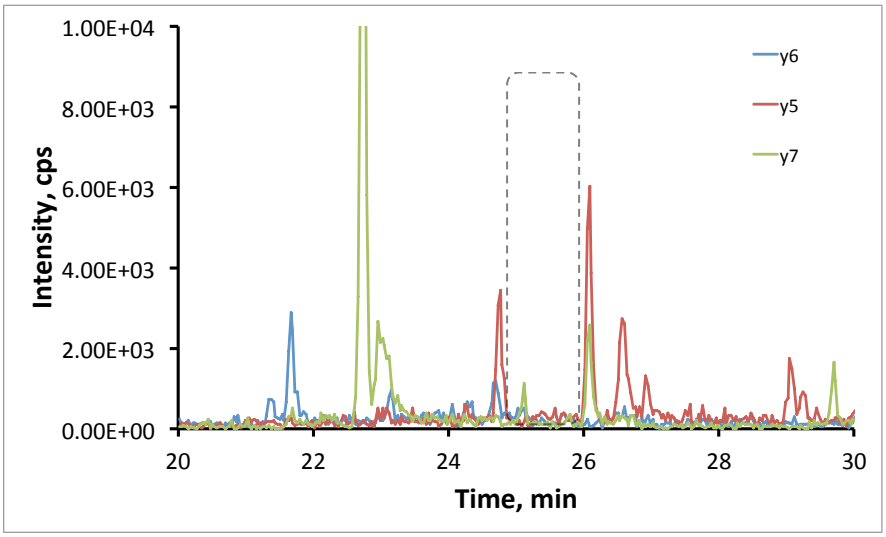


Figure 2.22 SRM-MS measurement of L-Arg-0 p50 peptide in Nfkb1$^{-/-}$ MEF.

Five µg of a tryptic digest of nfkb1$^{-/-}$ whole-cell murine embryonic fibroblasts measured by SRM-MS on a 5500 Q-Trap. The dashed box indicates the established elution window for the L-Arg-0 peptide VFETLEAR.

*MEF cell pellets provided by Dr. Buket Yilmaz.

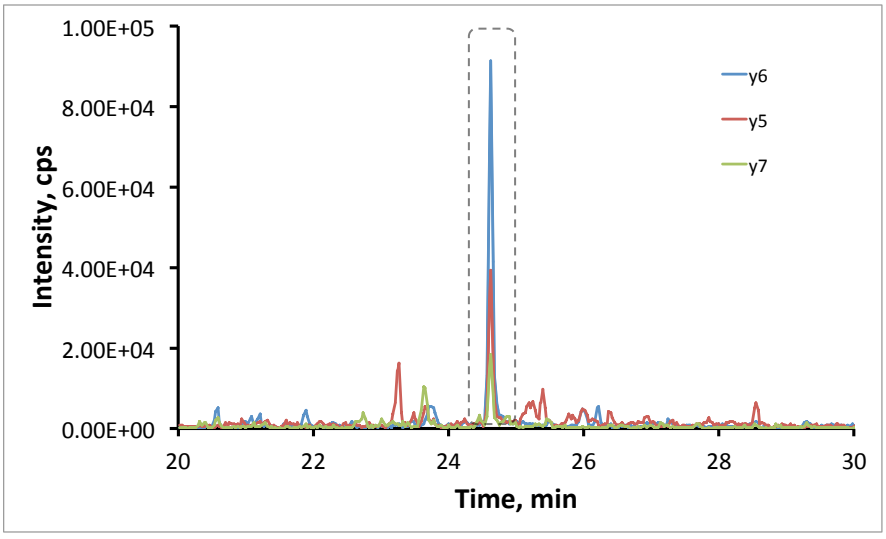


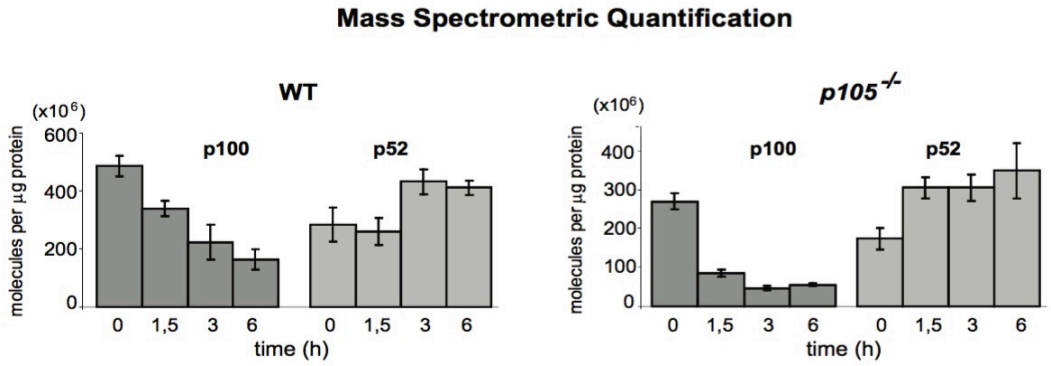
Figure 2.23 SRM-MS measurement of L-Arg-0 p50 peptide in wild-type MEF after Wessel-Fluegge extraction.

Five µg of a tryptic digest of Wessel-Fluegge-extracted wild-type whole-cell murine embryonic fibroblasts measured by SRM-MS on a 5500 Q-Trap. The dashed box indicates the SRM signals for L-Arg-0 VFETLEAR.

*MEF cell pellets provided by Dr. Buket Yilmaz.

The bar graph in Figure 2.24 shows the absolute mass spectrometric quantitation of p100 and p52 in MEF whole cell lysate following activation of the non-canonical NF-κB pathway by treatment with an anti-LT-βR antibody, 5G11b, as agonist [123]. Five micrograms of total protein was injected alongside 5 fmols of each H-Arg-10 labeled internal standard. The calculated peak area ratios underwent a correction for the contribution of p100/p105-derived p52/p50 peptides, via subtraction of the N-terminal peptide signals. SRM-MS analysis of stimulated wild-type MEF cells (Figure 2.24A, left) reveals a steady decrease in p100 with time, about 2.5 fold relative to the unstimulated sample after 6h, while p52 increases after 3h and is maintained at that level. An SRM-MS analysis of a similar LT-stimulation time course in $p105^{-/-}$ MEF cells (Figure 2.24A, right) supported the enhanced processing phenotype that was seen in the Western-blot analysis. After only 1.5 h p100 levels plummeted about 3-fold, reaching a lower endpoint than in the wild-type MEF. As expected, the sharper reduction in p100 is matched by an earlier and more pronounced increase in p52 levels.

A.



B.

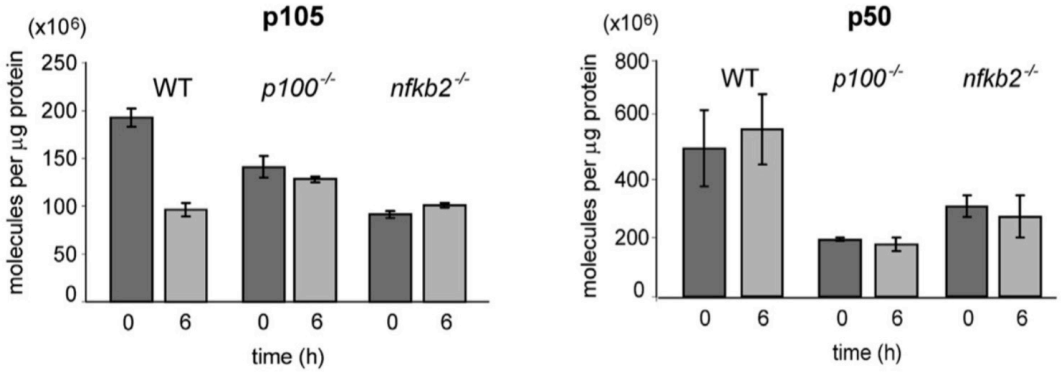


Figure 2.24 Absolute quantitation of p100, p52, p105 and p50 in MEF cells.

A. Absolute quantitation of p100 and p52 in wild-type (left) and p105 KO (right) MEF cells by SRM-MS.

B. Absolute quantitation of p105 and p50 in wild-type, p100 KO and nfkb2$^{-/-}$ MEF cells by SRM-MS.

MEF cells of the indicated type were subjected to LTβR-stimulated non-canonical NF-κB activation by treatment with the receptor agonist 5G11B. Total p100, p105, p52 and p50 were quantified by SRM-MS over three replicate injections, with the average of six transitions plotted as the number of protein molecules per microgram of protein input.

*All MEF cell culture and LT stimulation performed by Dr. Buket Yilmaz.

In the previously seen Western blot of LTβR-stimulated p105$^{-/-}$ MEF cells in Figure 1.2B it appears that the band for p100 has disappeared completely by the 3h mark, leading one to believe that the population of p100 has been entirely abolished in these cells. By SRM-MS, however, p100 remains visible, being readily detected even after 6h of stimulation. In comparison to anti-body-based detection of p100, the SRM-MS method is more sensitive.

3.1.7 Extended WT MEF stimulation.

The proof-of-concept experiments satisfied all three questions set forth, the established SRM-MS methods were shown to be selective and capable of replicating the trends previously seen in the non-canonical system using traditional molecular biology techniques. The ability to detect p100 after 6h of stimulation showed the methods to have a greater sensitivity compared to Western blotting. With this confidence longer duration LT-βR stimulation time courses were conducted, extended to 12h. Figure 2.25 shows the results for both the precursors (p100, p105) and the products (p52, p50). The MEF cell culture and LT stimulation was performed by Dr Buket Yilmaz. For each time-point, three injections were performed and two transitions quantified for each, resulting in the six calculated peak ratios displayed as black dots. These data illustrate the low deviation amongst the measurements and the SRM-MS method encouragingly was able to detect both p100 and p105 following the full 12h of stimulation, with the plot illustrating that the levels of these precursors remain suppressed and do not begin to recover for the duration of the time course.

Conversely, the products p52 and p50 increase in concentration until 4 to 6h and then maintain these levels for the remainder of the time. Both of these findings are consistent with the known “slow and persistent” dynamics of non-canonical NF-κB signalling, with similar kinetics governing the loss of precursors and the rise of products. Most interesting is the absolute quantitative information given in these plots. The basal levels of the products and precursors are revelatory, with p100 proving to be about 3-fold more abundant than p105. Both products are found at the 0 timepoint, likely caused by background proteasomal processing. Interestingly, a similar number of the two products are activated in response to the stimulus. Figure 2.26 presents the basal and 8h-stimulated concentrations of all of the proteins in nanomolar amounts, calculated by assuming an average MEF cell volume of 2 picolitres [124].

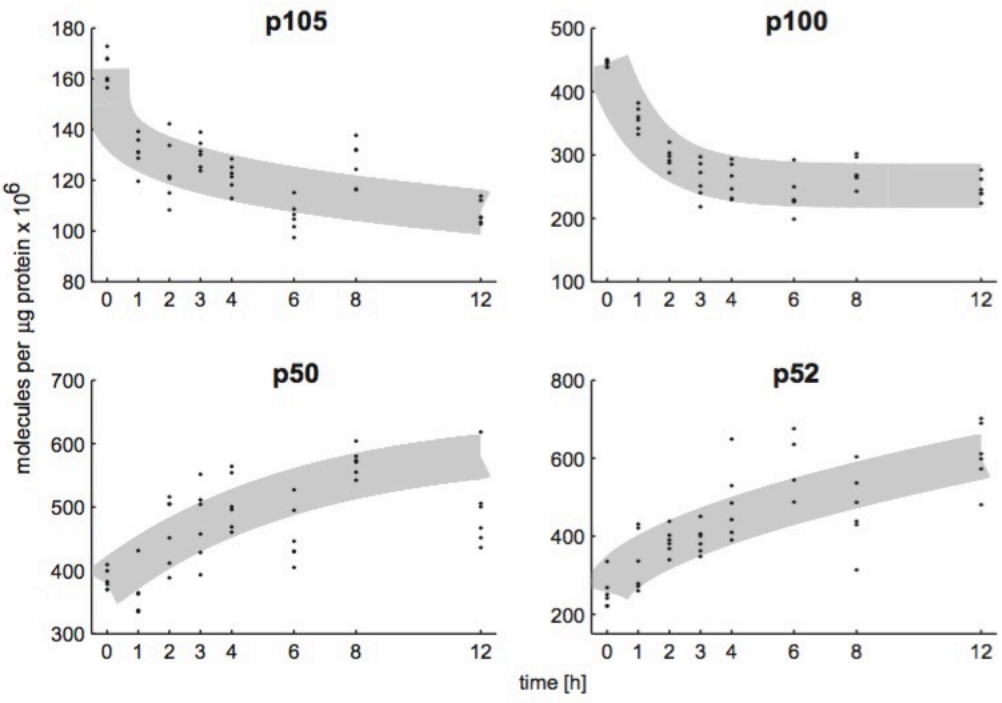


Figure 2.25 Absolute quantitation of p100, p52, p105 and p50 in LT-stimulated wild-type MEF cells by SRM-MS.

Wild-type MEF cells were stimulated with the LTβR-agonist antibody 5g11B to trigger non-canonical activation and cells were sampled at the indicated time points. Cell pellets were processed by in-solution digestion with trypsin and subjected to SRM-MS analysis to quantitate the levels of p100, p52, p105 and p50, normalized to the amount of total protein injected. Each sample was measured three times and the top two most-intense transitions used for the quantitation, resulting in six data points per time point, seen as black dots. The

trend curves are the best fit obtained through the data points by performing a least-square fit to a parameterized exponential, logarithmic or hyperbolic function curve.

*All MEF cell culture and LT stimulation performed by Dr. Buket Yilmaz.

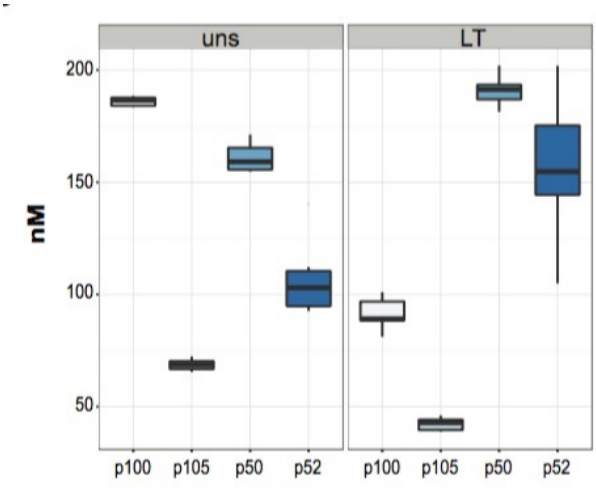


Figure 2.26 Basal and LTβR-stimulated concentration values from the absolute quantitation of NF-κB precursors and products in wild-type MEF cells.

Wild-type MEF cells were either unstimulated or stimulated with the LTβR-agonist 5g11B to trigger non-canonical activation and cells sampled after 8 h and subjected to SRM-MS absolute quantitation. A MEF cell volume of 2 pL was used to calculate nanomolar concentrations, and the box-plot shows the spread of the values from the six transitions quantified over three replicate injections.

*All MEF cell culture and LT stimulation performed by Dr. Buket Yilmaz.

3.1.8 Dynamic SILAC SRM-MS analysis of MEF LTβR stimulation time course reveals distinct precursor and product populations.

The extended LT-βR stimulation highlights the fact that this pathway remains “on” for many hours, with no indication of the precursors or products returning to their pre-stimulus levels even after 12h. One of the many transcriptional targets downstream of the pathway are the NF-κB components themselves as part of a negative feedback loop designed to restore equilibrium. Quantitative real-time polymerase chain reaction (qRT-PCR) data has shown that the NFKB2 gene is significantly up-regulated after about 6h[25]. The time scale of the response allows sufficient time for both gene transcription and subsequent protein translation, posing the question: what role do *de novo*-synthesized p100 and p105 play in the non-canonical signalling pathway?

SILAC is a quantitative proteomics approach well adapted to handle this query [75]. Cells grown in the presence of isotopically-coded amino acids will naturally incorporate them into proteins being synthesized. A special medium that has been deprived of lysine and arginine is supplemented with these molecules where ^{13}C and ^{15}N atoms have replaced ^{12}C and ^{14}N, respectively, creating a small difference in mass that is easily distinguished by a mass spectrometer. All of the peptides that function as proxies for their parent proteins have an arginine at the C-terminus (see Table 1), therefore it was sufficient to only use arginine composed of ^{13}C (Arg6). By replacing the Arg0 medium of the MEF cells at the time of stimulation with one with the Arg6 residues, new proteins synthesized subsequently will incorporate this form of the amino acid, while the proteins present prior to the stimulation will contain exclusively Arg0. Our naming convention refers to the "light isotope" species as L-Arg-0, while the "medium-heavy isotope" species is M-Arg-6. Spiking the "heavy isotope" synthetic H-Arg-10 internal standard peptide into the samples facilitates quantitation of both populations. This workflow is summarized in the schematic in Figure 2.27. Please note that MEF cell culture and LT stimulation was performed by Dr. Buket Yilmaz.

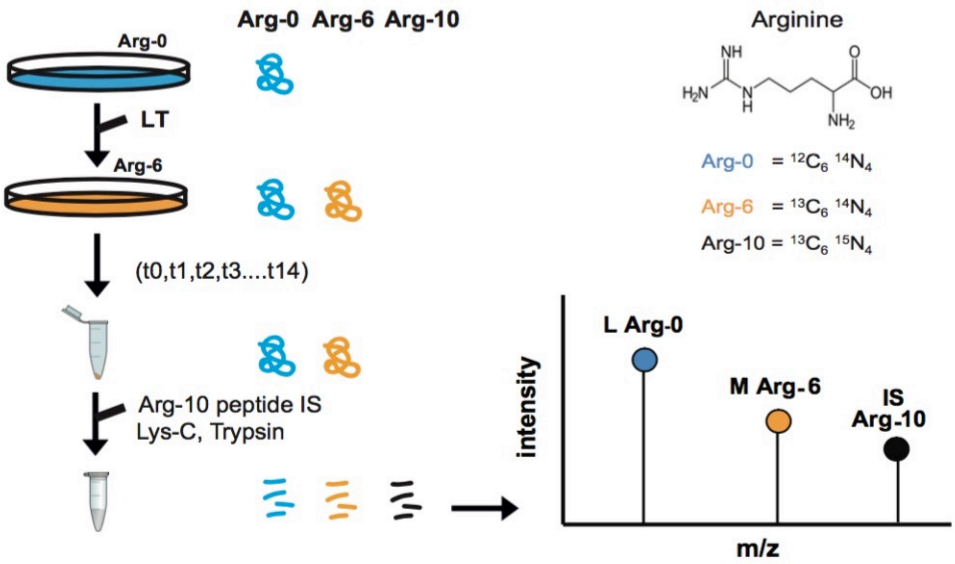


Figure 2.27 Workflow schematic of the Dynamic SILAC LT stimulation experiment.

Wild-type MEF cells were grown in light L-Arg-0 SILAC media and transferred to a medium-heavy M-Arg-6 SILAC media with or without accompanying LTβR stimulation. Cells were sampled at multiple time points and proteins quantified by SRM-MS using H-Arg-10 internal

standard peptides from JPT. The sample spectra indicates how the mass spectrometer can resolve peptide species with the same sequence but variable isotopic compositions.

The results of the Dynamic SILAC time course are presented in Figure 2.28 and show both the total protein levels as well as the contributions made by L-Arg-0 and M-Arg-6 populations in MEF cells undergoing LTβR stimulation. Examining the precursors p100 and p105, there is a decrease in the total protein levels (stacked bars) reaching a minimum after about 6 to 8 h. After 12 h, the original L-Arg-0 protein has been depleted to levels nearly below the detection limits of the assay. Newly-synthesized precursors (M-Arg-6) are evident within the first hour, and reach a plateau after about 6 h. Upon inspection of the products p52 and p50, it is apparent that the levels of the original protein present prior to stimulation remain nearly constant, while the increase in the total levels of the precursors and of both products comes from the *de novo* synthesized protein. Taking these observations together, it would appear that the original precursor proteins are preferentially degraded by the proteasome, to such an extent that they have been largely removed from the system by 12 h. Simultaneously, the *de novo* synthesized precursors appear to be preferentially processed by the proteasome, with the C-terminal half removed to yield the p52 or p50 products. Indeed, it is only the new protein material that results in an increase in the levels of the active transcription factors. This finding accounts for the late onset of the rise in the products, and therefore the transcriptional response to the stimulus itself, as new precursors are first synthesized and then processed. The Dynamic SILAC experiment has shown unequivocally that LTβR stimulation in MEF cells results in the precursors, at endogenous levels, are subjected to both complete degradation and specific proteolytic processing.

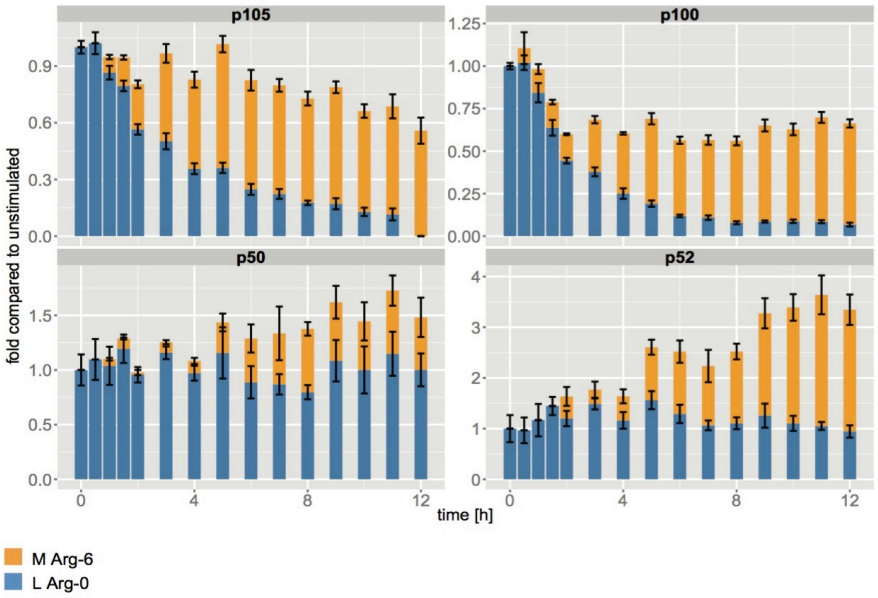


Figure 2.28 Quantitation of NF-κB precursors and products in the Dynamic SILAC LT-stimulation time course.

Wild-type MEF cells initially grown in L-Arg-0 SILAC media were moved into M-Arg-6 SILAC media containing the stimulation antibody 5G11b to trigger noncanonical activation. Cells were sampled at the indicated time points and processed for SRM-MS analysis in triplicate. L-Arg-0 (blue bars) and M-Arg-6 (orange bars) peptide contributions to total protein levels (stacked bars) of the precursors p100 and p105 and their products p52 and p50 are plotted, normalized to the unstimulated case at t=0.

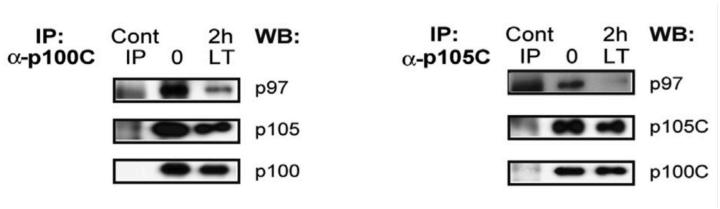
*The MEF cell culture and LT stimulation was performed by Dr Buket Yilmaz.

3.1.9 Inhibition of the ribosome, proteasome, or VCP-p97 blocks accumulation of *de novo* M-Arg-6 products.

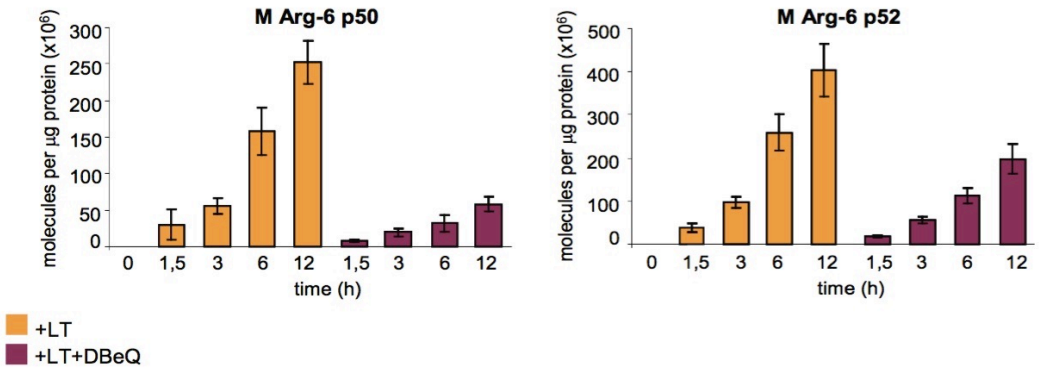
The partial proteasomal processing of the precursors p100 and p105 to give rise to the active transcriptional machinery is one of the most curious features of the non-canonical pathway. What mechanism exists to block complete proteolytic destruction at a precise site? The involvement of an ATPase, Ter94, in the regulated proteasomal processing of the Drosophila transcription factor Ci [27] was sufficiently similar to cast suspicion upon its mammalian counterpart, p97. Immunoprecipitation experiments (Figure 2.29A) revealed that p97 is associated with both precursors under resting and stimulated conditions in wild-type MEF cells, and with each cursor individually when the other was knocked out. A mutant p97 that had an inactive ATPase domain impaired the production of p52 relative to the wild-type in a NIK-overexpression system, in which p100 processing is constitutively “on” in the

absence of any upstream stimulation. The small-molecule inhibitor DBeQ [33] provides a convenient way of investigating the role p97 plays in the non-canonical pathway of a healthy wild-type MEF cell. Incorporating inhibition of p97 into the established Dynamic SILAC system used previously means the effect of p97 ATPase inhibition can be explored in the two populations of proteins, the pre-existing (L-Arg-0) and those synthesized *de novo* following stimulation (M-Arg-6). As evidenced in Figure 2.29B, the most pronounced effect was on the generation of M-Arg-6 products. DBeQ treatment dramatically reduced the accumulation of the M-Arg-6 p52 and p50 species, while having minimal effect on the L-Arg-0 p52 and p50 populations, which remained constant (Figure 2.29C).

A.



B.



C.

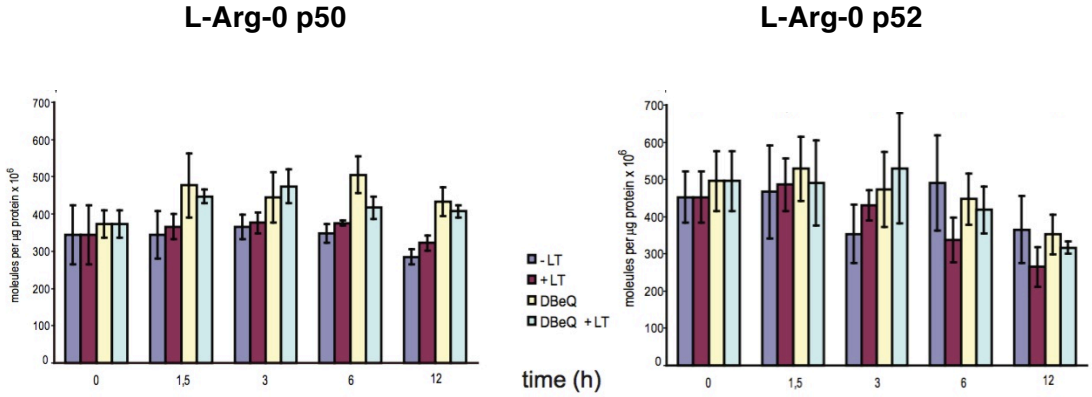
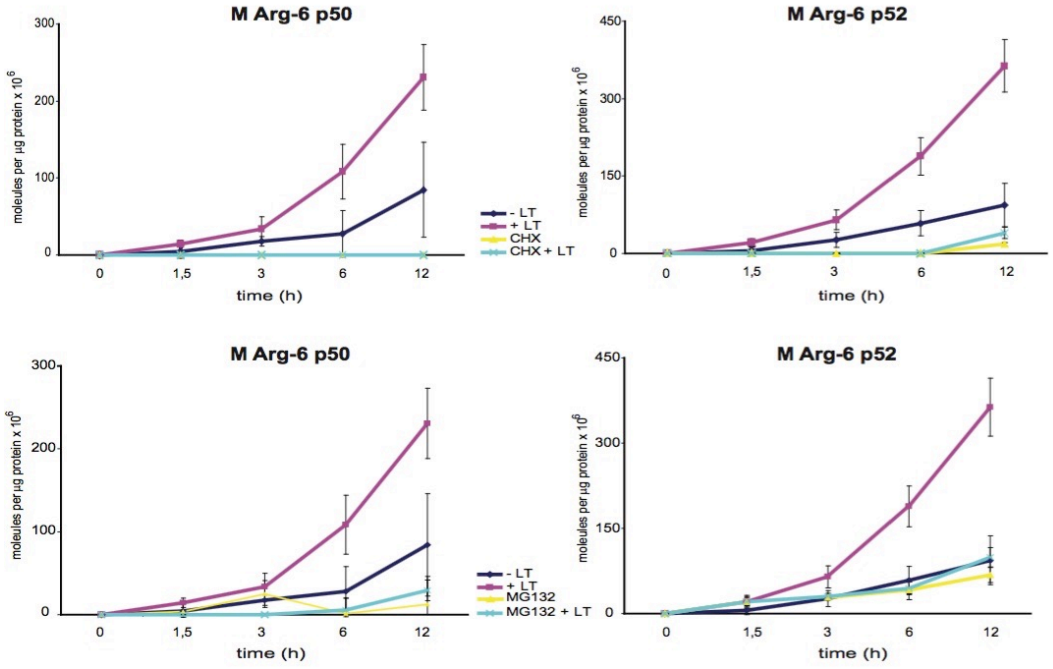


Figure 2.29 Dynamic SILAC SRM-MS evaluation of the influence of p97 ATPase activity on M-Arg-6 precursor processing.

A. Immunoprecipitation of either p100 and p105 copurifies p97, as detected by Western blotting [25]. Immunoblots courtesy of Dr. Buket Yilmaz.
B. Similar to Figure 2.28 above, wild-type MEF cells initially grown in L-Arg-0 SILAC media were stimulated by 5G11b in the presence or absence of the p97 inhibitor DbeQ after being moved to M-Arg-6 SILAC media and cells sampled at the indicated timepoints for absolute SRM-MS quantitation. Shown are the effects p97 inhibition has on the levels of M-Arg-6 products p52 and p50.
C. Wild-type MEF cells subjected to Dynamic SILAC as above with SRM-MS quantitation of L-Arg-0 product species, with or without LTβR stimulation, and with or without p97 inhibition with DBeQ.
*The MEF cell culture and LT stimulation was performed by Dr Buket Yilmaz.

The Dynamic SILAC methodology provides a way to study in detail effects of inhibition at other critical points of the non-canonical signalling pathway. Upstream protein synthesis required to generate NIK following TRAF-cIAP complex disruption can be blocked with cycloheximide (CHX), preventing tRNA translocation and nascent protein elongation at the ribosome. Figure 2.30 shows that in the context of precursor processing, ribosome inhibition essentially obliterates any generation of the M-Arg-6 p50 and p52 species, while the pre-existing L-Arg-0 products were largely unaffected. A similar outcome is attained via direct inhibition of the proteasome with MG132 treatment, an inhibitor blocking the hydrolytic activity, whereby M-Arg-6 p52 and p50 generation is blocked while no effect is exerted on the pre-existing L-Arg-0 p52 and p50 molecules.



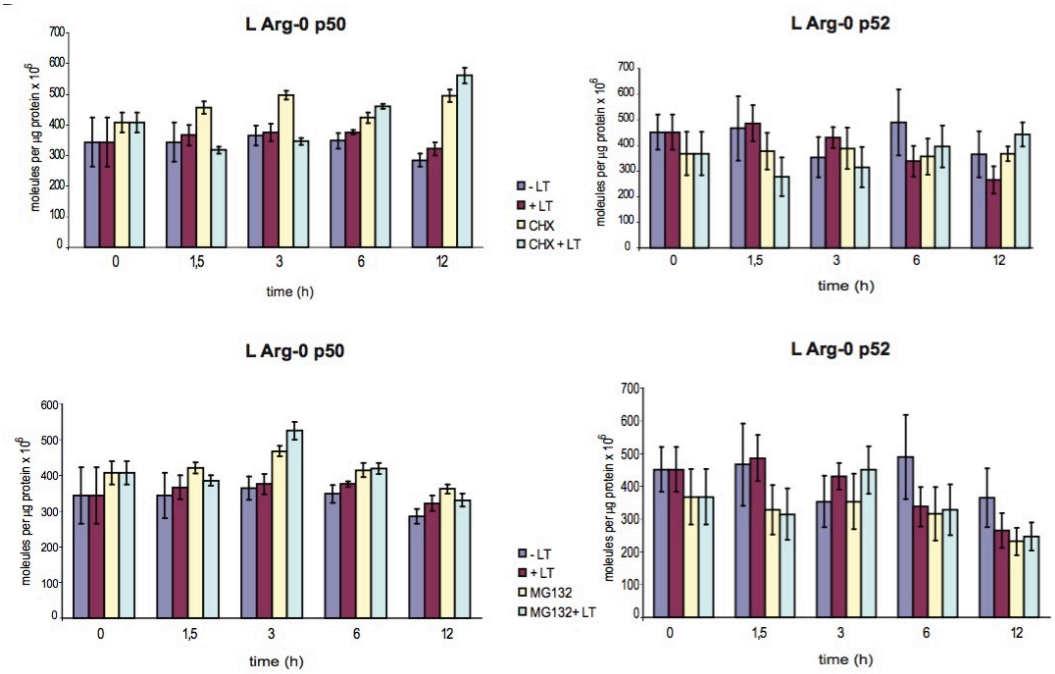


Figure 2.30 Effect of ribosome or proteasome inhibition on NF-κB precursor processing.

Wild-type MEF cells initially grown in L-Arg-0 SILAC media were left stimulated or unstimulated by LTβR agonist 5G11b in the presence or absence of either the ribosome inhibitor cycloheximide (CHX) or the proteasomal inhibitor MG-132 after being moved to M-Arg-6 SILAC media and cells sampled at the indicated timepoints for absolute SRM-MS quantitation. Shown are the effects CHX or MG-132 inhibition has on the levels of M-Arg-6 (upper 4 panels) or L-Arg-0 (lower 4 panels) products p52 and p50.

*The MEF cell culture and LT stimulation was performed by Dr Buket Yilmaz.

3.1.10 Mathematical modelling of the LT-dependent processing of the precursors predicts signal-dependent processing of a p100-p105 complex.

Compelling evidence exists that the NF-κB precursors p100 and p105 exist together in a complex and that they respond to an LTβR stimulation of the non-canonical response in an interdependent manner. One precursor is always found with the other in endogenous co-immunoprecipitation experiments and elute together in a high molecular weight fraction on a sucrose density gel. Most compelling are the processing phenotypes when one or the other precursor is eliminated: in a p100 knock-out, p105 fails to respond to the stimulus, while in a p105KO MEF, p100 responds more intensely, being processed with more rapid kinetics.

The absolute quantification of the products and precursors over a 12h LTβR stimulation time course and the resolution of the two protein populations represents a data set that forms the foundation for mathematical modelling efforts designed to gain insight into the signal-responsiveness of a proposed p100-p105 complex.

The SILAC SRM-MS analysis yielded the intriguing result that two distinct populations of precursor exist based on the way they responded to the stimulus. The blue section of the stacked bar graphs in Figure 25 indicate that the majority of the L-Arg-0 p100 and p105 are degraded completely, while the *de novo* M-Arg-6 p100 and p105 are diverted largely into the processing pathway, being converted to M-Arg-6 p52 and p50, respectively, by specific proteasomal hydrolysis.

Compiling these findings, mathematical models comprising free or unbound p100, p105, p52 and p50, in addition to the purported p100-p105 complex, were proposed and have been summarized schematically in Figure 2.31. The four models are differentiated by the outcome pathway stimulation has on the p100-p105 complex. Please note that the generation of mathematical models was done in collaboration with Dr. Bente Kofahl, Dr. Katherina Baum and Dr. Jana Wolf.

Model 1 (M1) The complex is unresponsive to the stimulus and forms a dead end that only dissociates into free p100 and p105.

Model 2 (M2) The p100-p105 complex is degraded in response to stimulus.

Model 3 (M3) The p100-p105 complex undergoes proteasomal processing to form the products p52 and p50 upon pathway stimulation.

Model 4 (M4) The p100-p105 complex is both processed and degraded in response to the stimulus.

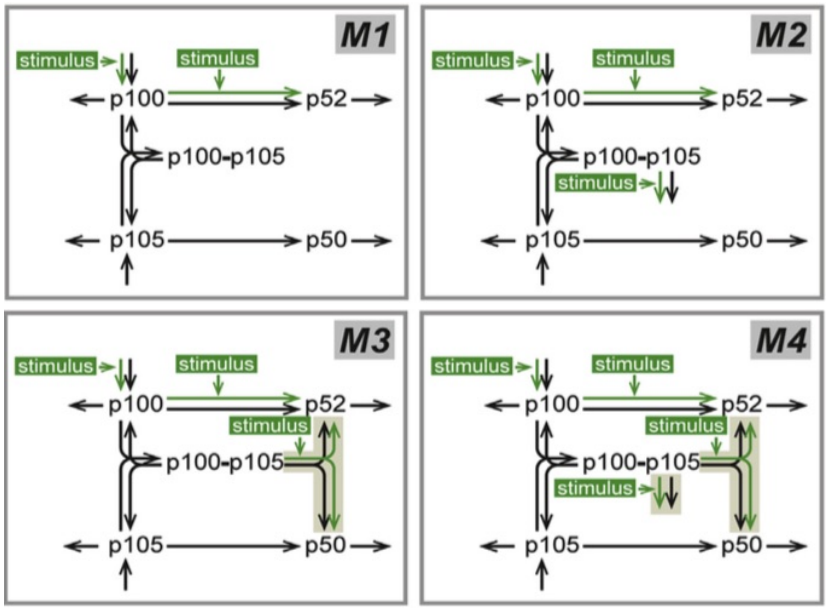


Figure 2.31 Proposed mathematical models of NF-κB precursor complex processing.

Proposed mathematical models describing the possible fates of a p100-p105 complex as well as the individual precursors and products. Included reactions are the *de novo* synthesis, degradation and processing of the precursors, either as independent entities or within a complex, and the degradation of products.

*Mathematical modeling performed by Dr. Bente Kofahl, Dr. Katherina Baum and Dr. Jana Wolf.

The processes shown in the schematic include precursor synthesis, degradation (turnover) and processing as singular entities, as well as formation of the complex and processing or degradation events acting on the complex as a whole. The model was intentionally designed to describe one aspect of the system, how a p100-p105 complex would respond to stimulation. Therefore the details and kinetics of the upstream events resulting in this stimulation have been omitted; for the purposes of this model the system is either “ON” or “OFF”. Likewise, events downstream of the processing/degradation of the individual precursors or complex have also been left out. Every arrow in the models is governed by a kinetic parameter, these parameters were estimated using an iterative computational process using the absolute concentrations found in the SRM-MS measurements over the duration of the time course, for both the stimulated and unstimulated condition. It should be noted that the SRM analysis cannot distinguish between free p100 and p105 and that bound

together in a complex, therefore these values are represented as being summed together: “p100 + p100-p105” and “p105 + p100-p105”.

After fitting the data to the four models and obtaining the estimates of the relevant kinetic parameters, a determination of their relative quality was calculated using the Akaike Information Criterion (AIC_c) {Akaike:1973wv}{Burnham:2004wc}. The ΔAICc value is a calculation of the differences in the AICc and is a useful metric for ranking the models in order of likelihood. Such a ranking is displayed in Figure 2.32, which indicates that model M3 is the most probable, at 83%, while M2 and M4 trail behind at 10% and 8% respectively. In model M3 the p100-p105 complex is processed to products in a stimulus-dependent manner. Model M2 (complex degraded in response to stimulus) and model M4 (complex both processed and degraded in response to stimulus) have been ranked as being significantly less likely but cannot be absolutely ruled out. The likelihood of model M1, with a completely unresponsive dead-end p100-p105 complex, has been entirely rejected by this calculation.

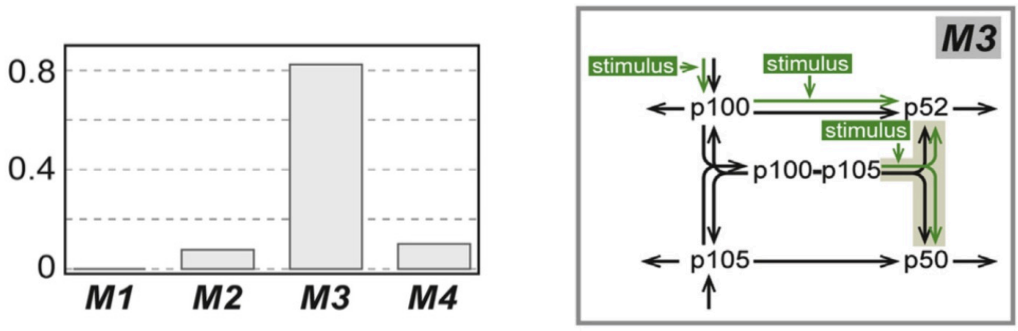


Figure 2.32 Calculated Akaike weights for Models M1 to M4.

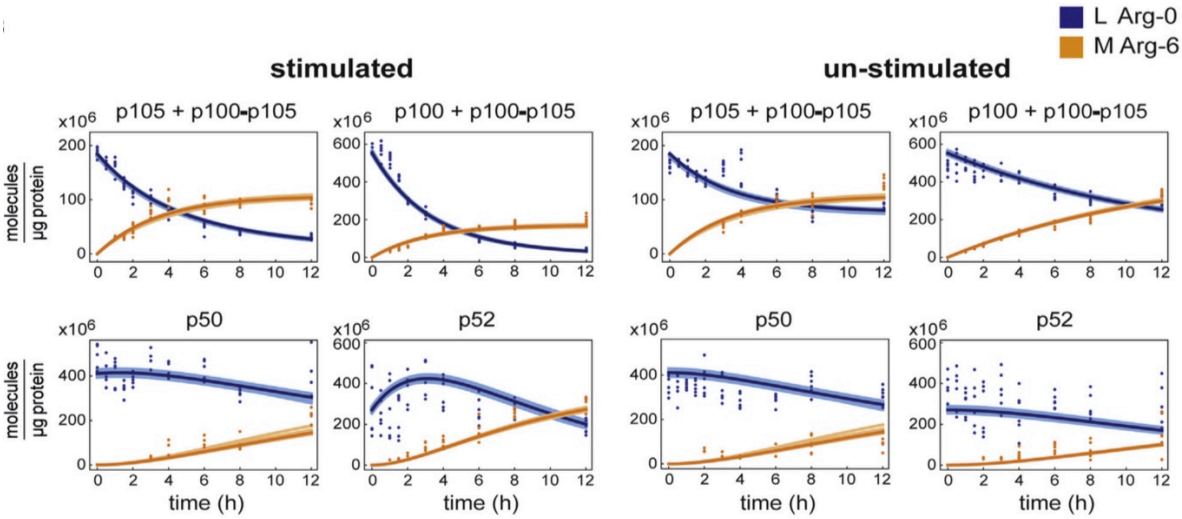
The higher Akaike weight indicates the more probable model. Model M3 predicts a stimulus responsive p100-p105 complex.

*Mathematical modeling performed in collaboration with Dr. Bente Kofahl, Dr. Katherina Baum and Dr. Jana Wolf.

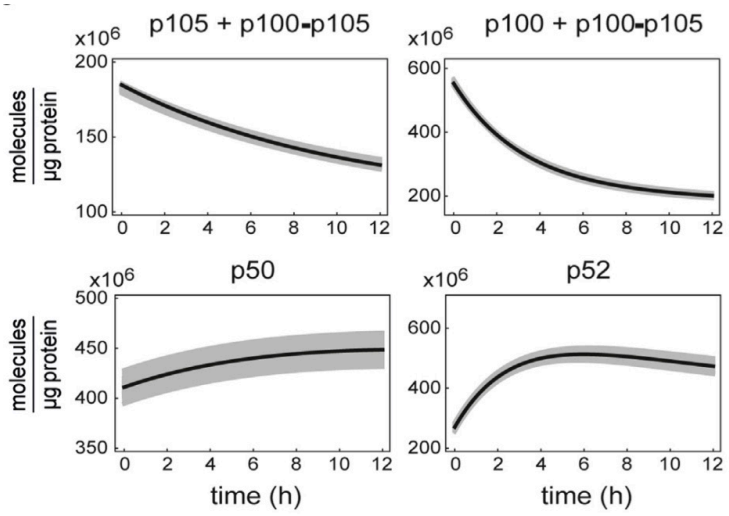
With a complete set of estimated kinetic parameters, the proposed models can now be used to perform simulations to predict the outcome. Figure 2.33A shows the simulations for the top ranked model M3 overlaid with the actual experimental values determined by the SILAC SRM-MS approach. Precursor and product are divided into their two populations, L-Arg-0 and M-Arg-6, and the simulated and observed values are in remarkably good agreement. Both of these populations have been summed to

give the total values in Figure 2.33B, with the results mirroring the established temporal changes in the p100, p105, p52 and p50 levels following the LTβR stimulation. A further form of validation is achieved by running simulations for model M3 in which important players are deleted to reflect the MEF knock-out systems tested. The p105 KO simulation (in which p105 levels are set to 0) replicated the altered p100 dynamics with respect to the wild-type that we have observed previously by Western blot and SRM-MS (Figure 2.33C). This can be rationalized by no p105 being present to 'absorb' its portion of the stimulus, resulting in proportionately more stimulus being available for the activation of p100, hence it being processed at a higher rate.

A.



B.



C.

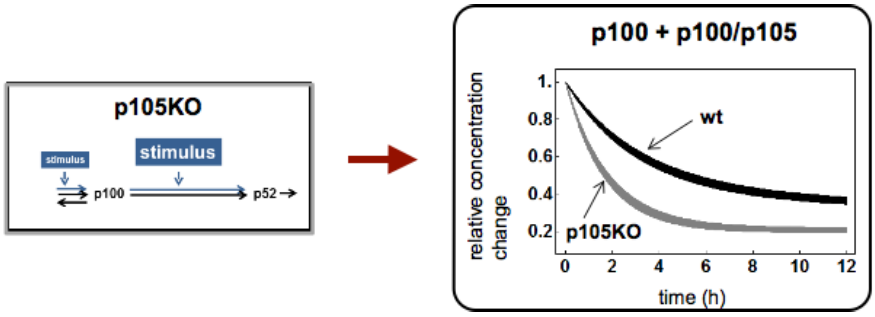


Figure 2.33 Model simulations of LTβR stimulation by the model M3.

A. Simulated time courses of stimulated or unstimulated MEF cells generated by Model M3 (solid lines) and overlaid with the experimental data (points), separated according to the protein species, L-Arg-0 in blue, *de novo* M-Arg-6 in orange. The 95% confidence intervals of the parameters are given by the shaded areas. SRM-MS analysis does not differentiate between individual precursors and precursor participating in complex formation, hence [p100 + p100-p105] and [p105 + p100-p105] represent the total contributions.

B. Simulated time course as given in A, but with the L-Arg-0 and M-Arg-6 protein levels summed to give the total amounts.

C. The p105 KO MEF phenotype is simulated by applying a proportionally larger stimulus to p100. Plotted versus a stimulation of wild-type MEF to demonstrate the enhanced kinetics of p100 processing.

*Mathematical modeling and simulations performed in collaboration with Dr. Bente Kofahl, Dr. Katherina Baum and Dr. Jana Wolf.

3.1.11 Expansion of SRM-MS analysis to the broader NF-κB pathway.

The SRM-MS analysis of the precursors p100 and p105 (and their respective products, p52 and p50) proved to be a worthwhile venture. Importantly, it mirrored the established trends previously observed using Western blot detection while exceeding its' limits of detection. The use of stable isotopes facilitated not only absolute quantitation, but through Dynamic SILAC also dissected the contributions made through *de novo* protein synthesis. Absolute values were critical in establishing a predictive mathematical model covering the mechanics of precursor complex formation and proteasomal processing.

Expansion of the SRM-MS technique to include additional players in the NF-κB signalling pathway would therefore be of great value in providing more quantitative detail than has been seen previously, and contribute to a mathematical model of broader scope comprising these additional players. The primary targets are the core NF-κB transcription factors. In addition to bolstering the methods for p100 and p105 by monitoring more peptides, peptides representing Rel A, Rel B and c-Rel will be included. As the IKK complex plays a central regulatory role in many signalling modes, SRM-MS detection methods for peptides from IKKα, IKKβ and IKKγ/NEMO would also be valuable. Additional targets include IκBα, GAPDH and Histone H2B. The inclusion of peptides for a cytoplasmic protein (GAPDH) and a nuclear protein (Histone H2B) is intended as a quality control for measurements in these cell fractions, to assess cross-contamination between the compartments. The histone peptide has another utility, as a signal for normalization of other protein concentrations in a similar manner as presented in Mann's Histone Ruler paper [125]. As the quantity of histone proteins (in nucleosomes) is fixed relative to the amount of genomic DNA, one can interpret their signal as a proxy for total cell number, a useful conversion in cases where counting cell pellets is difficult.

Prospective tryptic peptides were selected from the protein sequences according to the criteria presented earlier. Synthetic peptides (JPT Peptide Technologies, GmbH) were used to develop SRM-MS detection methods. After dissolving the peptides into the acetonitrile/ABC buffer, they were infused into the MS with the aid of a syringe pump and MS and MS/MS spectra acquired. These spectra established critical parameters, both the peptide's observed mass-to-charge (m/z) ratio and the fragmentation pattern following CID. Ramping of collision energies was again employed to select the optimal values for each transition to ensure maximum

sensitivity. Previous experience with optimization of the declustering potential proved these curves to be rather flat, so a single value of 130 V was chosen for all peptides in the method, maintaining constant conditions at the ion source for the entire analysis.

Table 2.2 includes the necessary SRM parameters for 26 (plus 2 oxidized forms of peptides) representing 12 proteins. Isotopically-heavy SpikeTide TQL peptides were ordered from JPT, with the C-terminal Lys and Arg residues substituted for $^{13}C_6{}^{15}N_2$ and $^{13}C_6{}^{15}N_4$ versions. In order to maximize applications for the methods, a peptide was chosen that had the same sequence in both human and mouse wherever possible, with mouse-only sequences indicated in parantheses beside the protein name. Where a sequence contains a cysteine residue, C[alk] indicates that the method assumes a quantitative conversion of the free thiol group to the alkylated form via reaction with 2-chloroacetamide. Methionine-containing peptides include two entries, one each for the unoxidized and oxidized (M**ox**) version, with Q1 and Q3 masses adjusted to account for the addition of a single oxygen atom (atomic mass 15.9994). This allows an estimate to be made of the contribution the oxidized peptides make to the population.

Q1	Q3	Protein/Peptide/z/Fragment ID	DP (V)	CE (eV)
561.3	963.5	NfkB2/ASDGSFSLPLK/+2/y9	130	20.9
561.3	848.5	NfkB2/ ASDGSFSLPLK.+2 y8	130	23.4
561.3	704.4	NfkB2/ASDGSFSLPLK/+2/y6	130	21.3
427.2	653.3	NfkB2/SLVDTYR/+2/ y5	130	17.4
427.2	554.3	NfkB2.SLVDTYR.+2/y4	130	19
427.2	439.2	NfkB2/SLVDTYR/+2 y3	130	29
402.7	691.4	NfkB2/LFGLAQR/+2/y6	130	22.7
402.7	544.3	NfkB2/LFGLAQR/+2/y5	130	22.1
402.7	487.3	NfkB2/LFGLAQR/+2/y4	130	22.7
466.3	631.4	NfkB1/NWATLAQK/+2/y6	130	19.2
466.3	560.3	NfkB1/NWATLAQK/+2/y5	130	22.2
466.3	459.3	NfkB1/NWATLAQK/+2/y4	130	21
663.4	961.6	NfkB1/NIHLHAHSLVGK/+2/y9	130	33
663.4	848.5	NfkB1/NIHLHAHSLVGK/+2/y8	130	34.2
663.4	711.4	NfkB1/NIHLHAHSLVGK/+2/y7	130	37
530.3	702.4	Rel A/DLEQAISQR/+2/y6	130	25.4

530.3	574.3	Rel A/DLEQAISQR/+2/y5	130	24.7
530.3	503.3	Rel A/DLEQAISQR/+2/y4	130	25.9
511.7	687.4	Rel A/GDYDLNAVR/+2/y6	130	21.4
511.7	572.4	Rel A/GDYDLNAVR/+2/y5	130	25.5
511.7	459.3	Rel A/GDYDLNAVR/+2/y4	130	26.9
522.8	931.4	Rel B/IC[alk]FQASYR/+2/y7	130	24.3
522.8	771.4	Rel B/IC[alk]FQASYR/+2/y6	130	24.4
522.8	624.3	Rel B/IC[alk]FQASYR/+2/y5	130	27.5
465.7	559.3	Rel B/EIEAAIER/+2/y5	130	21.5
465.7	488.3	Rel B/EIEAAIER/+2/y4	130	20.4
465.7	417.2	Rel B/EIEAAIER/+2/y3	130	19.6
793.9	947.4	c-Rel/LVSINLENPSC[alk]NAR/+2/y8	130	37.5
793.9	818.4	c-Rel/LVSINLENPSC[alk]NAR/+2/y7	130	37.1
793.9	704.3	c-Rel/LVSINLENPSC[alk]NAR/+2/y6	130	37.1
575.3	903.4	c-Rel/FVLNDWEAR/+2/y7	130	25.5
575.3	790.3	c-Rel/FVLNDWEAR/+2/y6	130	26
575.3	561.3	c-Rel/FVLNDWEAR/+2/y4	130	27.6
469.6	684.4	Traf6, mouse/HLQENTQLHMR/3+/y10	130	24.8
469.6	622.3	Traf6, mouse/HLQENTQLHMR/3+/b5	130	25.1
469.6	556.3	Traf6, mouse/HLQENTQLHMR/3+/y4	130	25.5
477.6	700.4	Traf6, mouse/HLQENTQLHM**ox**R/3+/y10	130	24.8
477.6	622.3	Traf6, mouse/HLQENTQLHM**ox**R/3+/b5	130	25.1
477.6	572.3	Traf6, mouse/HLQENTQLHM**ox**R/3+/y4	130	25.5
686.4	1044.6	Traf6, mouse/LTILDQSEALIR/2+/y9	130	35
686.4	816.5	Traf6, mouse/LTILDQSEALIR/2+/y7	130	35
686.4	931.5	Traf6, mouse/LTILDQSEALIR/2+/y8	130	35
567.3	905.4	Traf2/DLAMADLEQK/2+/y8	130	22.6
567.3	834.4	Traf2/DLAMADLEQK/2+/y7	130	23
567.3	703.4	Traf2/DLAMADLEQK/2+/y6	130	23.3
575.3	921.4	Traf2/DLAM**ox**ADLEQK/2+/y8	130	22.6
575.3	850.4	Traf2/DLAM**ox**ADLEQK/2+/y7	130	23
575.3	703.4	Traf2/DLAM**ox**ADLEQK/2+/y6	130	23.3
516.3	801.4	Traf2, mouse/ETFQDHVR/2+//y6	130	29.1
516.3	654.3	Traf2, mouse/ETFQDHVR/2+//y5	130	29.6
516.3	411.2	Traf2, mouse/ETFQDHVR/2+//y3	130	34.5
595.8	917.4	IKKγ/C[alk]LEENQELR/2+/y7	130	28.7
595.8	788.4	IKKγ/C[alk]LEENQELR/2+/y6	130	27.6

595.8	659.3	IKKγ/C[alk]LEENQELR/2+/y5	130	30.5
486.8	744.5	IKKα/VEVALSNIK/2+/y7	130	20.4
486.8	645.4	IKKα/VEVALSNIK/2+/y6	130	20.9
486.8	574.4	IKKα/VEVALSNIK/2+/y5	130	23
487.3	773.4	IKKα/SIQLDLER/2+/y6	130	20.6
487.3	645.4	IKKα/SIQLDLER/2+/y5	130	19.9
487.3	532.3	IKKα/SIQLDLER/2+/y4	130	25.2
545.8	977.5	IKKβ/LGTGGFGNVIR/2+/, y10	130	29
545.8	762.4	IKKβ/LGTGGFGNVIR/2+/, y7	130	29
545.8	705.4	IKKβ/LGTGGFGNVIR/2+/, y6	130	29.4
507.8	830.5	IKKβ/ALDDILNLK/2+/, y7	130	18.9
507.8	715.4	IKKβ/ALDDILNLK/2+/, y6	130	23
507.8	600.4	IKKβ/ALDDILNLK/2+/, y5	130	26.7
477.3	727.4	Histone H2B/LLLPGELAK/+2/y7	130	26
477.3	614.4	Histone H2B/LLLPGELAK/+2/y6	130	26
477.3	517.3	Histone H2B/LLLPGELAK/+2/y5	130	26
872.4	1126.5	Histone H2B/AMGIMNSFVNDIFER/+2/y9	130	43.4
872.4	1039.5	Histone H2B/AMGIMNSFVNDIFER/+2/y8	130	43.4
872.4	892.4	Histone H2B/AMGIMNSFVNDIFER/+2/y7	130	43.4
888.4	1126.5	Histone H2B/AM**ox**GIM**ox**NSFVNDIFER/+2/ y9	130	43.4
888.4	1039.5	Histone H2B/AM**ox**GIM**ox**NSFVNDIFER/+2 y8	130	43.4
888.4	892.4	Histone H2B/AM**ox**GIM**ox**NSFVNDIFER/+2/ y7	130	43.4
403.2	706.4	GAPDH/VGVNGFGR/+2/y7	130	22.7
403.2	649.3	GAPDH/VGVNGFGR/+2/y7	130	22.7
403.2	550.3	GAPDH/VGVNGFGR/+2/y7	130	22.7
435.3	770.4	GAPDH/VIPELNGK/+2/y7	130	24.1
435.3	657.3	GAPDH/VIPELNGK/+2/y7	130	24.1
435.3	560.3	GAPDH/VIPELNGK/+2/y7	130	24.1
585.7	1055.5	IκBα/DEEYEQSMK/+2/y8	130	30.7
585.7	926.4	IκBα/DEEYEQSMK/+2/y7	130	30.7
585.7	797.4	IκBα/DEEYEQSMK/+2/y6	130	30.7
593.8	1071.5	IκBα/DEEYEQSM**ox**K/+2/y8	130	30.7
593.8	942.4	IκBα/DEEYEQSM**ox**K/+2/y7	130	30.7
593.8	813.4	IκBα/DEEYEQSM**ox**K/+2/y6	130	30.7
495.2	852.4	IκBα/HDSGLDSMK/+2/y9	130	26.8
495.2	737.3	IκBα/HDSGLDSMK/+2/y8	130	26.8
495.2	650.3	IκBα/HDSGLDSMK/+2/y7	130	26.8

503.2	989.4	IκBα/HDSGLDSM**ox**K/+2/y9	130	26.8
503.2	989.4	IκBα/HDSGLDSM**ox**K/+2/y8	130	26.8
503.2	989.4	IκBα/HDSGLDSM**ox**K/+2/y7	130	26.8

Table 2.2 SRM method parameters for NF-κB pathway proteins.

3.2 The Protein-protein interaction network of NF-κB precursor p100.

SILAC SRM-MS analysis of the non-canonical NF-κB reponse led to the observation that two populations of the precursors p100 and p105, the pre-existing L-Arg-0 and the newly translated M-Arg-6, undergo distinct dynamics. The pre-existing L-Arg-0 population is preferentially degraded, with 90% of this population disappearing over the 12h time course, while the *de novo* synthesized M-Arg-6 precursors undergo a preferential processing process to transform them into the transcriptionally active p52 and p50 products. It was theorized that each form of precursor exists in a unique physicochemical context such that the cellular machinery can distinguish them and direct each down the appropriate enzymatic avenues. One possible scenario that might lead to such a distinction between the L-Arg-0 and M-Arg-6 population are differences in the p100-p105 complex, with each unique population being incorporated into a unique complex, whereby the set of interacting proteins are the distinguishing feature. Another potential scenario creating these differences is a chemical change to the precursor molecules themselves in the form of post-translational modifications. The discovery nature of this means that analytical approaches addressing these questions rely on data-dependent (shotgun) mass spectrometry, rather than targeted proteomics.

3.2.1 Determination of p100 interactome by immunoprecipitation.

Evidence abounds for an interaction between p100 and p105. On the physical side, co-immunoprecipitation of one precursor with the other indicates a close cellular proximity. The gel-filtration seen in Figure 1.2C shows both precursors migrating together in a high molecular weight complex, a complex that appears to shift to lower molecular weight upon stimulation with α-LTβR, suggesting a change in the complex's membership. Mechanistically, the contrasting p100 KO and p105 KO MEF phenotypes (Figures 1.2B & 2.24) show that p100 plays a crucial role; its absence abolishes the response, while the absence of p105 enhances the kinetics of p100 depletion. Mathematical modelling of this system has also supported the contention that p100 and p105 form a signal-responsive complex. Isolation of either precursor from the cell should be sufficient to enrich the other as well as complex components. The SRM-MS analysis of wild-type MEF cell lysate has determined the absolute levels of the precursors and products in basal and stimulated states; in both cases

p100 is shown to be in excess of p105, making p100 a better target for isolation by immunoprecipitation. The general approach to an MS-based interactome determination is to compare the identity and quantity of proteins between an experimental and a control pull-down, whereby normalization of the data reveals those proteins enriched in the pull-down of the target protein. In proteomics, two approaches for this normalization are utilized. Using a SILAC approach, one prepares cell lysates from differentially labeled cells and performs an experimental IP on one label type and a control IP on the other with an equivalent amount of input lysate. For example, cells labeled with light Lys0/Arg0 isotopes might be used for an experimental IP with a p100 antibody, while the heavy Lys8/Arg10 labeled cells are used for a control immunoprecipitation with IgG. A label-swap experiment in which the heavy-labeled cells are used for the experimental pull-down is also performed. Following elution from the beads, the pull-downs are combined and measured simultaneously in a single LC-MS run. The differential isotopic labels distinguish which proteins are derived from which pull-down, and each identified protein has an associated Intensity value for each label type, reflecting its abundance in the sample. The ratio of the Intensities of the two isotopic labels (i.e heavy/light or light/heavy) facilitates normalization and gives a measurement of relative enrichment between experiments. In contrast, the label-free quantitation (LFQ) approach does not utilize isotopic labelling of cells. Rather, equivalent pools of cell lysate are used for experimental and control pull-downs and the eluates measured separately by LC-MS. A minimum of three replicates of each condition are required for the MaxQuant LFQ algorithm to normalize the ratios. Each approach has its relative advantages. The SILAC methodology requires less instrument time as both conditions are measured simultaneously, and the isotopic labeling simplifies normalization. However sample complexity is effectively doubled by the isotopic labelling; identical heavy and light peptides will often both be sequenced, a redundancy that can limit the depth of protein identifications returned.

Please note that this project was in collaboration with Inbal Ipenberg of AG Scheidereit at the Max Delbrück Center, who performed all cell culture and immunoprecipitations. The SILAC methodology was chosen for initial experiments in p100 interactome determination due to its relative simplicity and the computational instability of early MaxQuant LFQ algorithms. Cells are grown in a special media supplemented with either L-Lys-0/L-Arg-0 or H-Lys-8/H-Arg-10 amino acids. Multiple passages are required to ensure the extent of incorporation of the heavy amino acids

approaches 100%. A portion of cells are lysed and the extracted protein digested with trypsin and measured by LC-MS as a quality control to determine the extent of labeling before proceeding with the planned immunoprecipitations. Figure 31 indicates an incorporation of 91.7%, sufficient to proceed with the SILAC experiment.

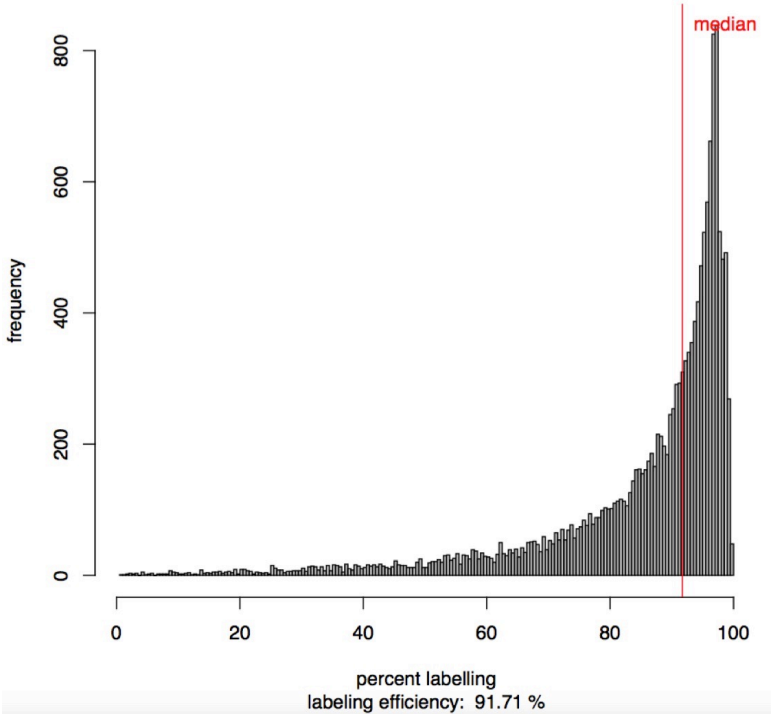


Figure 2.34 Incorporation of H-Arg-10 and H-Lys-8 residues in MEF cells by SILAC.
Twenty micrograms of MEF protein lysate was processed and digested with trypsin and the peptides measured by data-dependent analysis on a Q-Exactive mass spectrometers. Plotted is a histogram of the distribution of the H-Lys-8/H-Arg-10 in the detected tryptic peptides.
*MEF cell culture and isotopic labeling performed by Inbal Ipenberg.

Proteins bound to the p100 antibody were eluted with a chaotropic buffer containing 6M urea/2M thiourea and could be directly processed for LC-MS measurement, undergoing in-solution digest to convert them to tryptic peptides. Following a data-dependent shotgun analysis, the raw files were analyzed with MaxQuant. The intensities of proteins found in either the control (IgG, L-Lys-0, L-Arg-0 label) and experimental (anti-p100 H-Lys-8, HArg-10 label) pull-downs, measured in parallel in a single LC-MS run, were used to generate a Heavy/Light SILAC ratio. In total, 2953 protein groups were found in the MaxQuant analysis, what can be considered a high background. The Normalized Heavy/Light SILAC Ratios from MaxQuant Protein

Groups output file were converted to a Log2 scale and the MS1 intensities are represented on a Log 10 scale. Plotted together they create a so-called Christmas tree plot, found in Figure 32. The central cluster of proteins at a Log2 of 0 represents the majority, those proteins that have equal affinity to IgG as they do p100. The upper right hand portion of the plot is where the gene products enriched in the experimental immuoprecipitation can be found, in this case those proteins binding specifically to p100 and with an enrichment factor greater than 2 and an MS1 intensity greater than 1e9 counts per second (cps). The presence of p100 (NFkB2) at a high intensity and degree of enrichment is an indication that the immunoprecipitation was successful; absence of the bait protein would suggest problems with the experiment, such as poor antibody affinity or incomplete elution. In addition to the bait, all core members of the NF-κB transcription factor family have been enriched in the pull-down, including Rel B, p105 (nfkb1) and Rel A (nfkb3). The fifth member, c-Rel (mCG-8770), also appears in the plot, albeit with a lesser degree of enrichment. The ATPase p97/VCP previously implicated in the conversion of the precursors p100 and p105 into the products p52 and p50 was detected as well with a similar intensity as p100 (6e10 cps) but with a SILAC ratio of only 0.95, meaning it was not enriched in the p100 pull-down relative to the IgG control experiment.

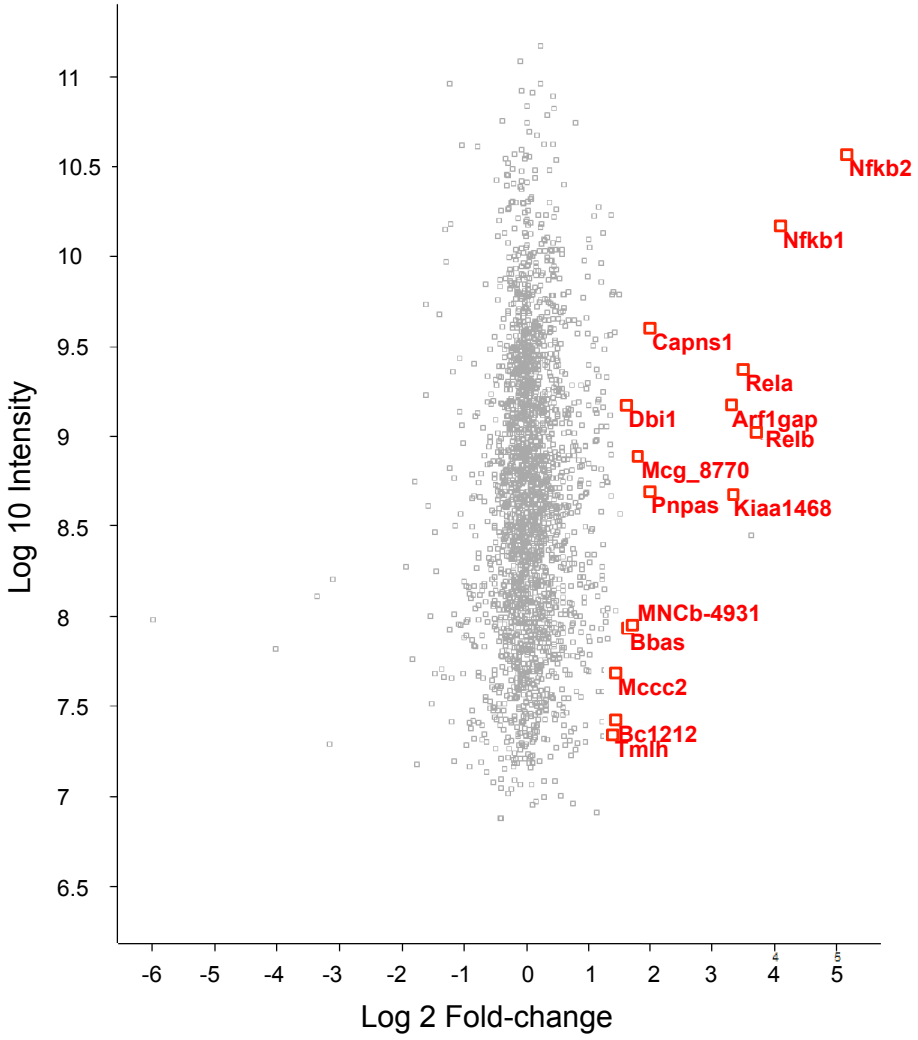


Figure 2.35 SILAC p100 interactome Christmas tree plot.

The eluted proteins in urea buffer that coimmunoprecipitated with p100 were subjected to an in-solution digest and the peptides measured in data-dependent mode on a Q-Exactive mass spectrometer. The data file was analyzed by MaxQuant 1.5.2.8 and the protein group file used to plot the Log10 transformed Intensities versus the Log2 transformed ratios. Enriched proteins with a Log2 fold-change greater than 1.5 are coloured in red with their gene name while background binders are indicated as gray circles.

*MEF cell culture and p100 immunoprecipitation performed by Inbal Ipenberg.

The lack of novel interactors found in the SILAC experiment prompted approaching the experiment using a LFQ approach. While this requires significantly more instrument time, measuring each experimental and control replicate separately, the absence of labels means the complexity of the analytes is reduced by half; each peptide exists only in the natural L-Lys-0/L-Arg-0 light format. Additionally, due to the

undersampling associated with shotgun measurements, repetitive analysis can augment the number of protein hits. The raw data for the label-free quantitation experiment was analyzed using the LFQ settings in MaxQuant 1.5.2.8. In total it returned 3030 protein groups, similar to the previous SILAC experiment. The log2 fold-change of proteins found in the control and experimental pull-downs (each measured in triplicate, six runs total) were plotted versus the –Log t-test p-value in Figure 33. Once again all transcription factors of the NF-κB pathway had a statistically significant enrichment in the p100 pull-down. The level of enrichment is comparable to the SILAC experiment presented above, with p100, p105, RelA and RelB clustering together and c-Rel (mCG_8770) enriched to a lesser extent. p97/VCP does not appear as a significant interactor of p100 under these conditions.

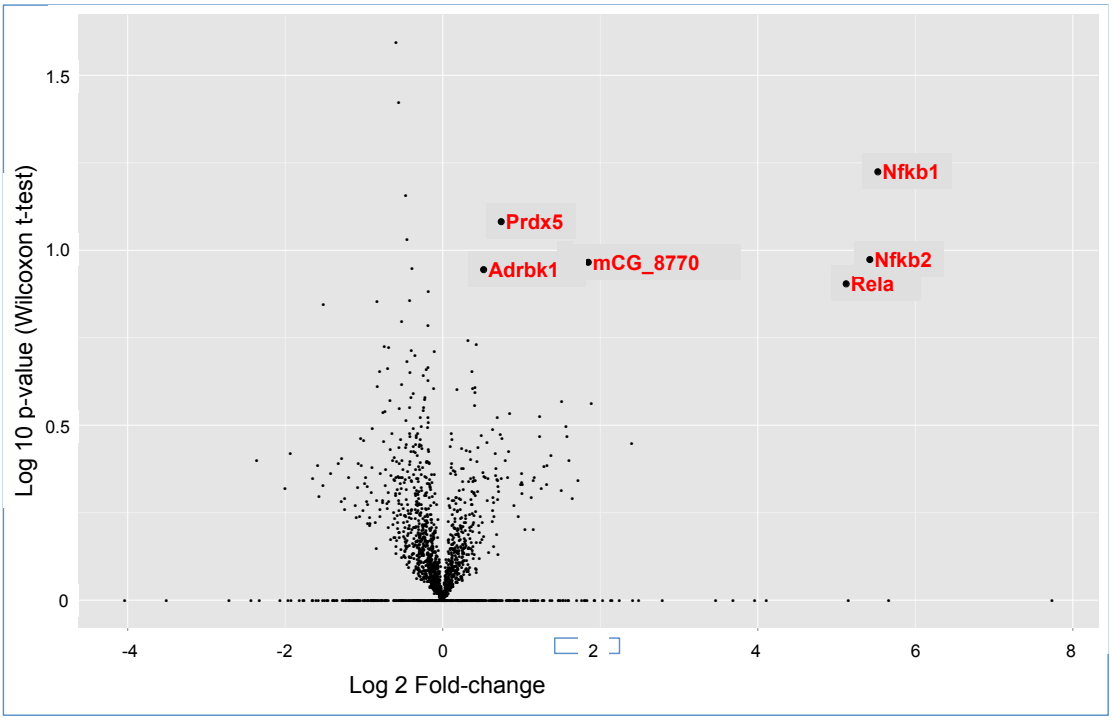


Figure 2.36 LFQ p100 interactome volcano plot.

The eluted proteins in urea buffer that coimmunoprecipitated with p100 were subjected to an in-solution digest and the peptides measured in data-dependent mode on a Q-Exactive mass spectrometer. The data file was analyzed by MaxQuant 1.5.2.8 with LFQ activated.

*MEF cell culture and p100 immunoprecipitation performed by Inbal Ipenberg.

3.3 MS screening for radiation-induced interactors of NEMO/IKKγ.

The classical immunoprecipitation approach to identify protein-protein interactors must be refined when shifting from Western blot to MS methods for the detection and analysis. While the advantages of MS over Western blotting in interactome studies are manifest, with hundreds of proteins routinely identified in eluates, care must be taken to avoid chemical additives common to molecular biology but having a deleterious effect in proteomics workflows. The tryptic digestion of proteins to form peptides and the conversion of these peptides into gas-phase ions in the MS ion source are critical junctures in the analysis that are susceptible to chemical interference. The search for protein interactors of IKKγ/NEMO, illustrates such a refinement of this approach. Please note that in this section all cell culture and immunoprecipitations were performed by Nadine Mikuda of AG Scheidereit at the Max Delbrück Center.

The non-canonical NF-κB activates gene transcription in response to many stimuli, among them genotoxic stress such as DNA double-strands breaks caused by exposure to radiation. It was previously demonstrated in the Scheidereit group that the response to DNA damage includes formation of a cytoplasmic complex involving NEMO that directs NF-κB activation [126]. The SILAC method was chosen as the basis for identifying inducible interactors of NEMO upon irradiation of cells in culture where NEMO is present at endogenous levels. Cells are labeled with light Lys0/Arg0 or H-Lys-8/H-Arg-10, with one label type exposed to radiation and the other not, as well as a label swap in a reverse experiment. Initial trial experiments were performed in MEF cells due to their ready availability in the lab with SILAC labeling already in place. The histogram in Figure 34 illustrates the incorporation of the H-Lys-8/H-Arg-10 labels into MEF proteins.

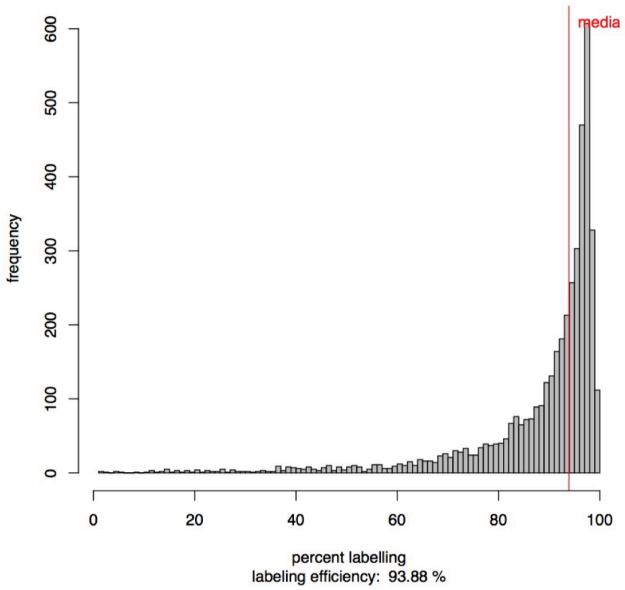


Figure 2.37 Incorporation of H-Arg-10 and H-Lys-8 residues into the MEF proteome.

Twenty micrograms of MEF protein lysate was processed and digested with trypsin and the peptides measured by data-dependent analysis on a Q-Exactive mass spectrometers. Plotted is a histogram of the distribution of the H-Lys-8/H-Arg-10 in the detected tryptic peptides.

*MEF cell culture and isotopic labeling performed by Nadine Mikuda.

3.3.1 Optimization of immunoprecipitation and MS conditions to improve enrichment of NEMO.

In the pilot experiments in MEF cells, protease inhibitors (Pefablock) were included in the IP wash buffer, a useful step to prevent background proteolysis that might alter the precipitated proteins and interfere with downstream antibody detection. However, as this eluate was prepared by an in-solution digest and analyzed by MS, carryover of these reagents into the subsequent digestion steps resulted in a poor preparation. The activity of trypsin was unsurprisingly inhibited, resulting in many missed cleavages (over 50% of peptides with one or more missed cleavages, 20% with two or more, compared to a typical overall rate of 20%). While NEMO was identified in this experiment, few of the other known interactors were found. After discussion of the results, it was decided that future IPs would omit Pefablock from the wash step. Additionally the cells would be fractionated and only the cytoplasmic portion used for the pull-down, as NEMO is partitioned in this compartment and including the complement of nuclear proteins would only serve to increase the background. Lastly,

incubation of the Dynabeads used with tRNA for 1h before use was designed to block non-specific binding to the beads. The next experiments with HepG2 cells were abandoned after the cells responded poorly to the SILAC culture media, forming large vacuoles and appearing unhealthy, raising concerns about whether such cells could be relied upon to respond normally to the radiation. However, analysis of these experiments was not entirely fruitless, revealing the detriment of another common additive in molecular biology and highlighting fundamental differences in sample preparation approaches between molecular biology and MS-based proteomics. NP-40, a polymeric detergent that aids in cellular fractionation by lysing outer plasma membranes but leaving the nuclear envelope intact, was included in one experiment and processed by an in-gel digest. A second equivalent experiment used douncing of the cells to shear the outer membrane and effect the fractionation. A comparison of the outcomes of MS analysis (Table 2.3) shows that the combination of NP-40 and in-gel digestion resulted in far fewer protein identifications (505) relative to douncing coupled to in-solution digest (1435). The number of NEMO peptides identified serves as a useful measure of the success of the immunoprecipitation. Additionally, the co-immunoprecipitation of known interactors of NEMO can be used to benchmark the experiments, as maintenance of these interactions is a promising indication that the IP and wash conditions were not unduly harsh.

Trial	Cell line	Cell lysis agent	Digestion protocol	Number of total protein groups	Number of IKKγ/ NEMO peptides	Percentage missed cleavages
P3	MEF	NP40	In solution	943	28	56.9*
P5	HepG2	NP40	In solution	1865	8	26.6
P6	HepG2	NP40	In gel	505	32	22.4
P7	HepG2	Douncing	In solution	1435	29	28.5
U1/U2	U2OS	Douncing	In solution	1822	20	33.5
U7/U8	U2OS	Douncing	In gel	1697	33	24.2

Table 2.3 Summary of IKKγ/NEMO IP optimization.
All samples measured on a Q-Exactive MS and raw data analyzed with MaxQuant 1.2.2.5. Asterisk (*) indicates that the protease inhibitor Pefablock was present during sample processing. Trial names are arbitrary.

Trial	IKK ϒ	IKK α	IKK β	Sharpin	HOIP	HOIL	NFκB1	NfκB2	Rel A	EDC4
P3	28	3	0	0	0	0	0	0	1	3
P5	8	1	2	0	3	3	8	5	3	9
P6	32	31	18	1	0	1	0	0	0	0
P7	29	1	1	0	2	0	0	3	4	4
U1/U2	21	3	5	1	1	0	4	3	3	5
U7/U8	33	23	23	0	4	0	14	15	10	23

Table 2.4 Number of peptides of related IKKγ/NEMO proteins found in MS screens[127].

The known interactors of NEMO had SILAC ratios around 1 in both the forward and reverse experiments, indicating their association with NEMO was not altered upon irradiation. A subset of the 1867 proteins identified in the U1/U2 screen had SILAC ratios greater than 2, indicating their association to NEMO was enhanced upon irradiation. Figure 2.38 represents these DNA-damage induced interaction partners of NEMO in a heatmap, with additional verification of the specificity of the interaction achieved using the protein frequency library (PFL) and the CRAPome database, comprised of 400 affinity-pull-down experiments, both of which aid in the correct categorization of known contaminants. The percentage listed refers to the occurrence of each interaction partner in these contaminant databases; therefore interaction partners with a lower percentage can be considered less likely to be associated non-specifically. The fourth column is an *in silico* prediction of which partners are likely to be targets of IKK phosphorylation, with more than half of the interactors returning a positive result.

Protein Name	PFL - Frequency of occurence (%)	CRAPome Ave SC	dbPTM predicted IKK sites
AF1Q	1.9	1.3	
AF4P12	2.9		+
AFCP	1.2		
AIP	3.9	1.4	
AKT2	4.6	2.1	*IKBKA*
AP3D1	17.8	2.8	+
ARG134	14.1		
ATXN10	nd	1.7	
BAX	2.2	1.2	
BCCIP	5.6	1.5	
BIRC6	2.9	1.7	
C10orf119	nd		
C14orf163	4.1		
C20orf37	1.5		
C7orf27	0.7		
CAND1	21.4	6.5	+
CAP43	4.6		
CIP2A	3.4		
CKAP5	20.4	9.4	+
CLH17	43.8		+
CLINT1	9.7	2.6	+
COBRA1	4.4		
COPB	22.4		
COPG2	16.3	15.4	+
CRMP1	10.7	2.1	+
DDB1	25.1	4.8	+
DDX20	9.7	2.6	+
DIS3	3.9	1.8	
EDC4	10.9	5.1	+
EDD	4.6		
EIF2B4	5.4	1.9	+
FACT140	18.5		
FACT80	17.8		+
FAM114A1	0.2	2	+
FAM91A1	5.1	1.9	+
FBXL1	nd		+
FOX2	3.4		
G22P2	46		
GALK1	2.9	3	+
GBF1	1.5	1.5	+
GCA	0.2	0	
GCN1L1	25.3	11.4	
GMPPA	0.7	1	
GOA	0.2		+
GTPBP1	4.4	1.7	
HCCS4	nd		+
HECTD1	1.2	2.4	+
HEATR2	1.7	1.3	
HSPC321	2.4		+
IPO8	13.4	21.7	+
KIAA0683	2.4		
KIAA0791	11.9		+
KLC2	12.9	3.1	+
MALT1	0.2	1	+
MAP2K2	1.7	1.7	+
MAP3K7IP1	10.9	5.5	+
MDA9	3.2		
METAP2	1	1.3	
MSTP011	1.7		
RPL9	40.6	5.7	
PAFAH1B2	3.4	2.6	
PCBP1	47.2	8.2	+
PI4KA	3.4	2.6	+
RNH1	7.8	19.3	+
RABGGTA	1.7	1	+
RFC2	13.6	2.4	
ROCK1	6.8	3.6	+
SAMD4B	0	0	
SMAP1	10.7	2.1	+
SPAG2	0.7		
SRP14	19.5	2.6	
TSG101	1.2	1	
TUBB2	90.5		+
UBA5	0.5	1.5	
YWHAE	45.3	9.4	
Zw10	7.8	1.6	+

Figure 2.38 Radiation-inducible interactors of NEMO.

Interactors of NEMO enriched in forward (Lys8/Arg10 labeled cells irradiated) and reverse experiments. Frequency of occurrence in databases of non-specific interactors found in 400 affinity pull-down screens analyzed by LC-MS. Interactors predicted to be substrates of the IKK complex.

*Courtesy of Nadine Mikuda.

The inducible interactors were also subjected to a Gene Ontology analysis (Figure 2.39) to identify the biological processes most enriched. As expected, “Response to DNA damage stimulus” is near the top of the list, but it was the presence of “mRNA metabolic processes” which turned attention on to enhancer of mRNA-decapping protein 4 (EDC4), which is involved in the degradation of mRNA. Further investigation revealed that upon irradiation, IKK binds and phosphorylates EDC4. This leads to an increased formation of p-bodies, cytosolic structures that are the sites of mRNA processing, as well as recruitment of mRNA decapping enzymes Dcp1a and Dcp2. Ultimately, this work was able to show that the stability of a significant number of transcripts is regulated by the IKK and EDC4 signaling axis, ascribing a new role to IKK in protein regulation that extends beyond transcriptional regulation.

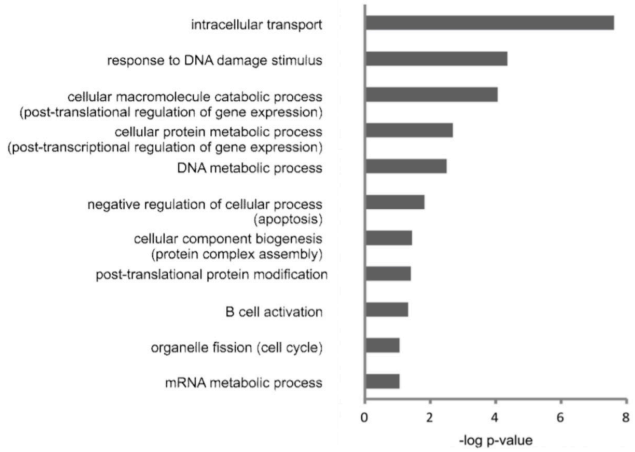


Figure 2.39 Gene ontology analysis of the inducible interactors of NEMO.

*Courtesy of Nadine Mikuda.

3.4 Analyzing polyubiquitination by mass spectrometry.

3.4.1 Ubiquitin linkage profiling by SRM-MS.

The mode of polyubiquitination on a substrate is a critical feature in determining the downstream fate of that protein. Primary ubiquitination sites on a target protein can be identified using mass spectrometry by looking for the lysine residues with a diglycl moiety (K-GG), appearing in a MS/MS spectrum as a lysine augmented by 114.03 Da. Ubiquitin's internal lysines are modified by diglycyl moieties by the process of polyubiquitination, with seven lysines (K6, K11, K27, K29, K33, K48 and K63) and the N-terminus (M1) possessing the primary amine functionality required for conjugation. Upon digestion with trypsin, eight unique peptides are formed, each one specific to a particular linkage type. SRM-MS methods for these peptides were published by Mirzaei *et al* [128] and adapted for use in our laboratory, with the exception of the M1/linear peptide which was developed in-house for these studies. Table 2.5 lists the transitions for the detection of endogenous ubiquitin linkage-specific peptides; transitions for heavy-labeled peptides were also published but are not listed here. Analysis of a whole cell lysate in the presence of a known amount of heavy internal standard can show global levels of each linkage type across the proteome. However, it is more informative to assess the polyubiquitination chain type conjugated to a single substrate, and therefore an immunoprecipitation to enrich this particular target is the starting point for further analysis in each of the studies presented below.

Q1	Q3	Ubiquitin linkage type; Sequence	DP (V)	CE (eV)
460.6	305.2	K6, MQIFV[K_GG]TLTGK	130	27
460.6	519.3	K6, MQIFV[K_GG]TLTGK	130	24
801.4	1002.5	K11, TLTG[K_GG]TITLEVEPSDTIENVK	130	38
801.4	1131.6	K11, TLTG[K_GG]TITLEVEPSDTIENVK	130	38
701.0	944.0	K27, TITLEVEPSDTIENV[K_GG]AK	130	28
701.0	802.4	K27, TITLEVEPSDTIENV[K_GG]AK	130	36
408.7	503.3	K29, A[K_GG]IQDK	130	28
272.8	390.2	K29, A[K_GG]IQDK	130	15
546.6	643.3	K33, IQD[K_GG]EGIPPDQQR	130	36
546.6	740.4	K33, IQD[K_GG]EGIPPDQQR	130	29

487.6	617.8	K48, LIFAG[K_GG]QLEDGR	130	19
487.6	347.2	K48, LIFAG[K_GG]QLEDGR	130	33
748.7	1015.5	K63, TLSDYNIQ[K_GG]ESTLHLVLR	130	32
748.7	1067.6	K63, TLSDYNIQ[K_GG]ESTLHLVLR	130	46
520.3	740.4	Ub internal, EGIPPDQQR	130	29
520.3	643.3	Ub internal, EGIPPDQQR	130	30
440.2	506.3	M1/Linear, GGMQIFVK	130	27
440.2	634.4	M1/Linear, GGMQIFVK	130	27
793.4	978.5	K11 yeast, TLTG[K_GG]TITLEVESSDTIDNVK	130	40
793.4	891.4	K11 yeast, TLTG[K_GG]TITLEVESSDTIDNVK	130	40
698.4	940.0	K27 yeast, TITLEVESSDTIDNV[K_GG]SK	130	29
698.4	804.4	K27 , TITLEVESSDTIDNV[K_GG]SK	130	36
416.7	503.3	K29 yeast, S[K_GG]IQDK	130	25
416.7	262.1	K29 yeast, S[K_GG]IQDK	130	25
896.0	802.4	actin (ACTB_HUMAN), SYELPDGQVITIGNER, y7	130	36
896.0	1086.6	actin (ACTB_HUMAN), SYELPDGQVITIGNER, y10	130	47

Table 2.5 Polyubiquitin linkage specific SRM methods adapted from literature.

The linear/M1 linkage-specific peptide was not used in the Mirzaei *et al*, Method parameters were developed manually on an ABI Sciex 5500 Q-Trap using the approach detailed for the precursors and products of NF-κB.

3.4.2 Polyubiquitin and parkin.

3.4.2.1 Parkin promotes linear polyubiquitination of NEMO by LUBAC.

Parkin is an E3 ubiquitin ligase linked to Parkinson's disease with a role in neuroprotective pathways. *In vitro*, parkin prevents cell death in response to mitochondrial and endoplasmic reticulum stress, while *in vivo* mutations in the parkin gene resulting in a loss of function cause autosomal-recessive parkinsonism. With regard to mitochondrial stress, parkin has been linked to a pathway that removes damaged mitochondria by mitophagy [129], and it was theorized that this role is the foundation for parkin's prosurvival function. Please note that this project was in collaboration with Dr. Anne-Kathrin Müller-Rischart of AG Winklhofer at the Ludwig Maximilians University, who performed all cell culture and immunoprecipitations. In MEF cells where parkin was knocked-out, Dr. Müller-Rischart observed a significant increase in apoptotic cell death, implicating the mitophagy pathway [130]. Systematically knocking out other components of the mitophagy pathway, PINK1, ATG5 and p63, did not hinder parkin's ability to block apoptosis, but removal of the ubiquitin-like (UBL) domain from parkin's N-terminus did. These findings indicated that parkin's roles in mitophagy and stress protection are regulated by different pathways. Parkin is known as a stimulator of the canonical NF-κB pathway [131], itself linked to anti-apoptotic activity, and therefore a possible link between them was pursued. As NEMO is an essential mediator of the canonical NF-κB response, MEF cells with NEMO knocked out are deficient in responding to canonical stimuli. It was shown that in these cells mitophagy was unaffected but parkin's ability to block apoptosis was prevented, and that this activity is restored upon rescuing the cells by expressing NEMO by transient transfection. Parkin is an E3 ligase and the NF-κB pathway is highly dependent on ubiquination, both degradative and non-degradative, in order to propagate signals, including linear ubiquination of NEMO [103]. The only other protein known to undergo linear ubiquitination is RIP1 [132], a kinase also associated to the NF-κB pathway. Linear polyubiquitin chain formation is catalyzed by the linear ubiquitin assembly complex (LUBAC), comprised of the E3 ligases HOIL-1L and HOIP and Sharpin as an adaptor protein [133] [134]. A link between parkin and LUBAC was therefore investigated using SRM-MS as a tool to quantitate changes in the various polyubiquitin linkage types. This was facilitated by using the UBAN domain of NEMO, known to have affinity for linear ubiquitin chains and can

therefore be exploited to enrich them [135].

HEK293T cells were transfected with the LUBAC components HOIL-1L and HOIP, as well as a wild-type parkin or a mutant form lacking the UBL domain. Lysates of these cells were incubated with a recombinant UBAN domain with a Strep tag to enrich for linear ubiquitin chains, which were then relatively quantified by SRM-MS using heavy-labeled analogues of the internal polyubiquitin linkage peptides listed in Table 2.5. Figure 2.40A illustrates that the expression of wild-type parkin in these cells is associated with a three-fold-increase in linear ubiquitination, while this activity is severely compromised when parkin's UBL domain is removed. The effect that endogenous parkin has on linear ubiquitination was investigated by knocking-down parkin levels with siRNA in the same cell system (HEK293T), and Figure 2.40B shows that suppression of endogenous parkin expression is associated with an approximately three-fold decrease in linear ubiquitination.

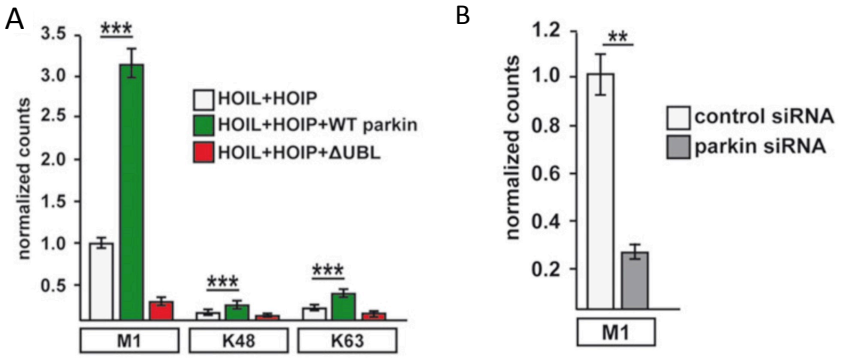


Figure 2.40 Parkin promotes the linear ubiquitination of NEMO.

A. SRM-MS relative quantitation of linear (M1), K48- and K63-linked polyubiquitin following affinity-purification of HEK293T lysates with Strep-linked UBAN. Samples measured in triplicate with top two transitions used for quantitation with standard deviation indicated.

B. Relative quantitation of linear polyubiquitin in HEK293T cells treated with siRNA to knock down parkin. Measured by SRM-MS as in A.

*HEK293T cell culture and immunoprecipitation performed by Dr. Anne-Kathrin Müller-Rischart.

3.4.2.2 Parkin is a component of LUBAC and binds preferentially to HOIP.

Parkin's effect on linear ubiquitination was further compounded by showing that it physically interacts with the LUBAC complex. Two cell lines of HEK293T cells were made overexpressing the LUBAC components, one with HOIP, HOIL-1L-HA and parkin, the other with HOIP-HA, HOIL-1L and parkin, with the HA tag used to enrich these proteins and their interactors. After analysis of the two pull-downs by shotgun mass spectrometry following an in-gel digestion, the data was analyzed by MaxQuant. Figure 2.41 reveals that Parkin binds preferentially to HOIP. Strikingly, it does so to an even greater extent than the known LUBAC component Sharpin, itself also exhibiting a preferential binding to HOIP. Other known interactors of HOIP and HOIL-1L (TRIM25, PSMD2, PSMD7, PSMC4) appear in the central cluster and do not favour binding to one LUBAC component over the other.

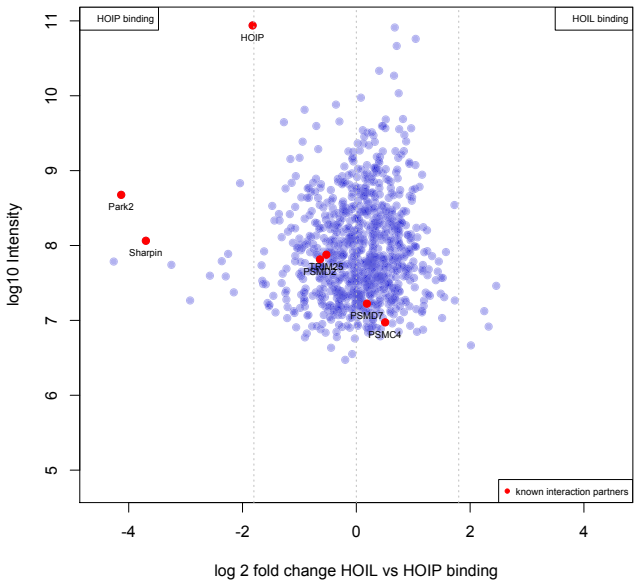


Figure 2.41 Parkin is a component of LUBAC and binds preferentially to HOIP.

HEK293T cells were transfected with HOIP-HA, HOIP-1L and parkin, or HOIP, HOIL-1L-HA and parkin and subjected to immunoprecipitation with α-HA antibody. Proteins bound to HOIP or HOIL-1L were measured and quantified on a Q-Exactive MS and the Log2 ratio of HOIP and HOIL-1L plotted versus the Log10 transformed normalized MS intensity. Known interactors of the bait proteins are coloured in red.

*HEK293T cell culture and immunoprecipitation performed by Dr. Anne-Kathrin Müller-Rischart.

3.4.3 Ubiquitin and protein kinase A signalling.

3.4.3.1 ARHGAP36 induces K63-linked polyubiquitination of PKAC.

Cyclic adenosine monophosphoate (cAMP), a derivative of adenosine monophosphate, acts as an intracellular signal transduction agent and is involved in several biological processes, acting as a second messenger downstream of G-protein coupled receptors. One such function is the activation of protein kinase A (PKA), a serine/threonine kinase with roles in regulating diverse activities such as development, proliferation and metabolism [136]. PKA is composed of a dimer of regulatory subunits (PKAR) each bound to a catalytic subunit (PKAC), occupying its active-site cleft and inhibiting its function. Upon binding of cAMP to PKAR, PKAC dissociates and is rendered active, facilitating phosphorylation of PKA targets. Tight regulation of this activity is achieved through multiple means, including by small protein kinase inhibitor proteins (PKI) that act as pseudosubstrates blocking the kinase site [137], or a mechanism by which PKAR is subjected to proteasomal degradation in a cAMP-dependent manner [138]. A ubiquitin-dependent degradation of PKAC was also suspected as this is a typical pathway for turning off protein kinase activity.

Please note that this following project was carried out in collaboration with Dr. Rebecca Eccles of AG Rocks at the Max Delbrück Center. In a previous MS screen for interactors of PKAC, a Rho-GTPase-activating protein family member ARHGAP36 was identified. Further co-immunoprecipitation studies confirmed the interaction, and it was demonstrated that ARHGAP36 inhibits the activity of PKAC by binding to its active site cleft via a pseudosubstrate motif. Beyond simple inhibition to downregulate the kinase activity, it was observed in HEK293T cells that expression of ARHGAP36 coincides with depletion of PKAC, as measured by Western blotting, evidence pointing towards a degradation pathway. A SRM-MS analysis of linkage-specific polyubiquitin peptides was performed, measuring their levels in a PKAC immunoprecipitation in HEK293T cells transfected with PKAC-YFP in the presence or absence of co-expressed Flag-ARHGAP36. Figure 2.42 demonstrates that their coexpression leads to an accumulation of K63-linked polyubiquitin with no change in levels of the K48-linkages classically associated with protein degradation. While inhibition of the proteasome by epoxomicin had no effect on PKAC levels, inhibition of endolysosomal acidification by bafilomycin partially restored PKAC.

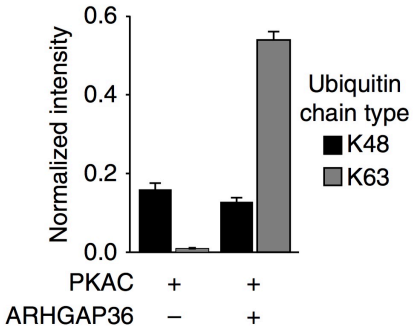


Figure 2.42 Coexpression of PKAC and ARHGAP36 results in increased levels of K63-linked polyubiquitination.

HEK293T cells were transfected with PKAC-YFP with or without FLAG-ARHGAP36 and a immunoprecipitation performed with an α-GFP antibody. Eluates were processed and SRM-MS analysis was used to quantify the relative levels of K48 and K63-linked polyubiquitination after normalizing protein loading based on PKAC levels determined by data-dependent measurement of a portion of sample on a Q-Exactive MS. SRM-MS Samples injected twice and the top two transitions used for quantitation and standard deviation calculated.

*Cell culture and immunoprecipitations performed by Dr. Rebecca Eccles.

3.4.3.2 Intensity-based quantification of neuroglioblastoma proteins.

A shotgun MS analysis of a neuroglioblastoma (NGP) whole cell lysate was quantified by a method known as Intensity-based absolute quantification (IBAQ) [139]. In this approach, the overall MS intensity of each protein detected is normalized by dividing this intensity by the number of theoretically observable tryptic peptides each protein can yield, a value converted to an absolute via calibration with a spiked in protein standard mix. In the absence of this calibration. this method enables a relative quantitation of all of the proteins found in the data set, ranking them according to their abundancy in the sample. In Figure 2.43, a histogram shows the distribution of relative abundance for all of the measured proteins (n = 4448). ARHGAP36 and PKAC appear in the same bin, indicating that they have a similar abundance. The accompanying bar plot shows the $\log_{10}$ abundance of several proteins relative to ARHGAP36; PKAC is at 0.95 fold, and therefore essentially equimolar to ARHGAP36. These findings confirm that ARHGAP36 is a viable regulator of PKA kinase activity. On a short-term time scale, endogenous and equimolar levels of ARHGAP36 block the kinase activity with its PKAC

pseudosubstrate motif, while on a longer-term time scale it promotes its K63-polyubiquitin mediated endolysosomal degradation.

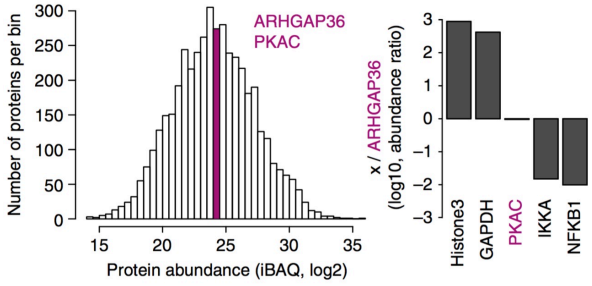


Figure 2.43 Intensity-based quantification shows ARHGAP36 and PKAC found in equimolar amounts in neuroglioblastoma cells.
Data-dependent MS analysis of whole-cell neuroblastoma lysate subjected to an IBAQ quantitation to sort proteins by relative abundance. Histogram (right) indicates the abundance distribution of all measured proteins (n = 4448) with the bin that ARHGAP36 and PKAC co-occur in coloured pink. PKAC is the sum of the two subunits of PKAC, PRKACA and PRKACB. Bar plot (left) indicates the abundance of the proteins relative to ARHGAP36. PKAC has a 0.95 fold abundance relative to ARHGAP36.
*Cell culture performed by Dr. Rebecca Eccles.

4 Discussion

4.1 Inception and evolution of proteomics.

The term “proteome”, first coined in 1994 [140], referred to the entire observable protein complement of a cell. This was a time when MS-based proteomics was still in its infancy, indeed before the word proteomics was even used, and the ultimate goals must have seemed lofty and idealistic. Although the identification of a handful of proteins might take months of work, the awesome potential of the nascent technology had clearly been recognized for its ability to transform protein science. Ten years later Mann and Tyers put forth an updated definition, describing proteomics as “almost everything post-genomic: the study of the proteome in any given cell and the set of all protein isoforms and modifications, the interactions between them, the structural description of proteins and their higher-order complexes” [141]. This expanded definition reflected the massive leaps forward that had already occurred in MS-based proteomics, driven by innovations in technology and informatics, and adopted in droves by researchers persuaded at the sheer power of MS-based proteomics approaches.

Ten years further onward and the once-lofty goals are beginning to become routine, with complete proteomes of organisms starting to come within grasp. Starting with lower-order organisms [142] all the way up to human proteomes [143] [144], what was once a fantasy can now be achieved readily, often in as little as a day. This exponential rise in the power of proteomics parallels that of the genomic field [145]. The Human Genome Project, a massive global effort to sequence the three billion nucleotide base pairs of our genetic code, took 13 years and cost $3 billion, reaching completion in 2003 [146]. Thanks to next-generation sequencing technology, the same goal could be achieved in days for a tiny fraction of the cost. In proteomics the development of highly-sensitive instrumentation and robust methodologies has propelled widespread growth in the field, offering immense advantages over the direct and indirect protein identification and characterization methods that came before. Foremost of these is the specificity with which proteins are identified by the detection and sequencing of multiple peptide proxies. Over the last 20 years, the basic workflows of proteomics have become increasingly standardized, with similar sample preparation and instrumental analysis methods adopted by groups around the globe. In most shotgun or data-dependent analyses, the measurements proceed in an unbiased way; the

researcher does not have control over what peptides are being selected for sequencing. Generic workflows and common instrumentation makes the sharing of raw data in public repositories an aid in research projects worldwide [147].

With the advent of deep-proteome sequencing and approaching full proteome coverage, one surprising observation has been just how similar the proteomes of different tissues are. Rather than being differentiated on the basis of the types of proteins expressed, in practice there are very few proteins that are unique to a given tissue [148]. This pervasive proteome expression was confirmed by the Human Protein Atlas, a global antibody-based effort to map protein expression and produce validated anti-bodies for each protein-coding gene [149]. Differentiation of these tissues must therefore be realized by differential protein abundances and arrangements, how many molecules there are and with what other molecules they interact. Hence, the ability of MS to precisely quantify proteins emerged as a vital property, as only through quantitation will we characterize these differences and reach a deeper understanding. Quantitative proteomics strategies involving the use of stable-isotope coded peptides and proteins dominate for their ability to correct for a number of experimental and instrumental variations, though label-free approaches have also steadily gained traction.

In the context of a discovery shotgun experiment, the instrument operates in an unbiased manner, sampling whatever peptide ions are most abundant at any given time during the chromatographic elution for MS/MS sequencing. While many thousands of peptides can be identified this way, there are also inherent problems. First, due to its data-dependent nature and the size of the digested proteome, the instrument samples a slightly different subset of this population each run, leading to a level of irreproducibility even between identical samples. Second, sampling only the most abundant ions means the less abundant signals are neglected, resulting in a systematic underrepresentation of these analytes in data sets [112] [110]. This is particularly true for the transcription factor class of proteins, which tend to be expressed at low levels in the cell [150] [151] [152]. How, then, to reliably measure them?

4.2 Targeted proteomics for analytically challenging scarce protein species.

Targeted proteomics strategies address both of these problems. Focusing the MS analysis on a pre-determined list of peptide analytes means they will be reproducibly sampled each time, even when other more abundant species are co-eluting from the chromatographic column. While not generating the same volume of data as discovery proteomics, targeted methods allow researchers to answer the specific questions they are interested in. SRM-MS on a triple-quadrupole MS is the most mature method in the targeted proteomics toolbox, and the method of choice for reliable and robust quantitative measurements across many samples and even between different labs and instrument platforms. SRM-MS has long been a “gold standard” for small-molecule drug candidate quantitation [153], measurements that need to be precise and reproducible if they are to satisfy powerful regulatory bodies like the Food and Drug Administration (FDA) in the U.S.A. Continuous development and refinement in QQQ technology brought the sensitivity of these instruments to a level matching the needs of proteomics. Early examples of SRM on plasma proteins [154] and Streptococcal virulence factors [155] demonstrated its feasibility, but it was the massive studies targeting the complete yeast proteome undertaken by the group of Rudi Aebersold, that demonstrated the power of the technique to the proteomics community. Picotti *et al* detected and quantified yeast proteins across the entire abundance range with a success rate of 90% [156]. Following this breakthrough, and spurred on by the concurrent creation of databases to share SRM methods, other groups started to apply this technique and publish their work, including groups interested in quantifying transcription factors such as the NF-κB family. Xiang *et al* quantified all NF-κBs, with the exception of c-Rel, as part of a project to characterize a model system of drug resistance in multiple myeloma, after first fractionating proteins using SDS-PAGE [157]. Zhao and colleagues measured activated Rel A molecules present in nuclear fractions following enrichment using a single-stranded DNA aptamer that binds free Rel A with high affinity [158]. In a different publication, the same group measured protein markers of the innate immune response in cellular fractions [159]. Simicovic *et al* probed the levels of key transcription factors during the terminal phase of adipogenesis [160]. What the above studies have in common is employing a fractionation or enrichment scheme in order to first reduce the complexity of the measured proteome, and as such should refrain from making claims of absolute quantitation. In two cases, cell lysates were resolved by SDS-PAGE before being

subjected to an in-gel digestion to convert embedded proteins into tryptic peptides, with heavy peptide internal standards spiked in later. Provided this SDS-PAGE procedure involves completely efficient processes to get the proteins into the polyacrylamide matrix and a lossless extraction of the resultant peptides, this could be true. However, the systematic efforts by the group of Schevchenko using isotope-coded peptides has shown that peptide recovery from the gel varies from 13 to 40% and is both protein- and peptide-dependent [121]. As it is not feasible to measure the extraction efficiency for all peptides in this way, the quantitation of in-gel preparations should be considered as relative, not absolute. Likewise, for affinity enrichment of a target protein, the efficiency of the pull-out is highly variable and sub-stoichiometric, so a claim of absolute quantitation would require a precise characterization of this pull-down process for each sample to define what percentage of the total population was captured.

In the present work, SRM-MS methods were designed and optimized to perform absolute quantitation of the NF-κB precursor molecules, p100 and p105, and their respective products, p52 and p50, following an in-solution digest of a protein solution derived from lysis of intact MEF cells. Such a preparation represents a maximum of sample complexity in proteomics, a level of complexity that is not fully defined [109]. While the number of protein-coding genes an organism possesses can be approximated, the number of discrete protein products actually present in a cell is complicated by alternate splice forms and post-translational modifications, a limitless landscape of possibilities further amplified by the ever-changing dynamics of biology. Faced with such a significant analytical challenge, taking measures to reduce sample complexity via fractionation or affinity enrichment are understandable steps to improve the chances of detecting one's target in the downstream MS measurement. However, such manipulations are difficult to quantify for the above-mentioned reasons, and introduce unknowns into the process that make absolute quantitation a challenge unless carefully controlled and monitored. In addition, fractionation or affinity enrichment is generally not reproducible across many samples and would therefore introduce further uncertainties. The goal here was a method that could probe hundreds of samples generated from time course experiments; after activating the non-canonical NF-κB pathway by treatment with a lymphotoxin B receptor agonist and sampling cells over a 6-12 h time window. In previous studies this detection was accomplished via Western blot, the standard technique for protein detection in biology labs, as evidenced by over 54,000 citations of the original 1979

publication [161]. Here the authors noted in the abstract that the transfer of proteins from gels to nitrocellulose was not quantitative. Coupled with antibodies of dubious specificity and non-linear detection properties, a Western blot is not the ideal tool for making precise quantitative measurements.

4.3 Critical facets of the SRM-MS method development process.

The selection of tryptic peptides on which to build an SRM method is a crucial step as it will ultimately determine the sensitivity of the assay. The signal response from tryptic peptides derived from the same protein can differ by 100-fold [162], a fact that could make the difference between a confident detection or a signal indistinguishable from the background noise. Unfortunately it is very difficult to predict any given peptide's MS performance based on the sequence alone. Some computational tools do exist, trained by existing data resources [163] [164] [165], with some incorporating up to 500 properties in their scoring algorithm [118]. The non-cohesiveness of results among platforms pushes some in the direction of public data repositories, such as PeptideAtlas [166] [167] and SRMAtlas [113], to simply adapt the methodologies already used by others. At the time of method development for this project, no information about NF-κB peptides was available, and indeed the majority of entries in these databases arose from yeast as a result of the work of Picotti *et al*. Data from shotgun experiments can also be used to identify peptides that are reproducibly detected, but these data generally come from non-QQQ instrument platforms, where the mode of fragmentation results in differing fragmentation patterns [168]. In this study, the in-house availability of a solid-phase peptide synthesizer provided the means acquire MS/MS spectra on the Sciex 5500 Q-Trap, the QQQ platform to be used.

The first step in the development of an SRM-MS assay is the characterization of the peptide's MS traits, such as the precursor m/z and charge state, as well as the MS/MS fragmentation pattern. Even in the presence of synthetic by-products (and contaminants), steady infusion of the synthetic peptides directly into the MS via electrospray source to give a stable signal made it possible to confirm the success of the synthesis by the presence of the expected m/z and to determine the charge state of the peptide under experimental conditions. While this is generally a straightforward calculation for a small peptide, the Enhanced Resolution scan of the peptides in Figs. 2.4 and 2.9 makes visualization of the spacing of the ^{13}C isotopes simple. When comparing the two types of precursor scans, EMS and ER (Figs 2.3, 2.8 vs. 2.4, 2.9),

one observes the trade-off for the higher resolution in terms of intensity, with a loss of nearly 10-fold. In SRM mode, the instrument allows one to set a resolution on the Q1 and Q3 m/z signals, filtering the selected ion through a window 0.35 (High), 0.7 (Unit) or 1.0 (Low) m/z units wide. A narrow bandwidth (high) on Q1 allows more precision in precursor ion selection, but the reduced ion transmission means less sensitivity for that target and therefore must be considered carefully. It is also noteworthy that a wider selection bandwidth offers a higher tolerance for drift in the calibration of the quadrupole, drift that might otherwise reduce ion transmission by sampling an m/z value off-centered from the maximum. For this study, the resolution was set to Unit for both Q1 and Q3.

The product ion or MS/MS spectra (Figs 2.5 and 2.10) are the foundation of any SRM-MS method development. Here the collision-induced dissociation (CID) fragmentation pattern is revealed, making selection of the most abundant fragments straightforward. Also valuable is the relative fragment abundance, as the same hierarchy of signals will be present in the nested peaks characteristic of SRM-MS data. Referring to Fig. 2.14, one sees the ranking y7 > y6 ≈ y8 seen in the MS/MS spectrum replicated in the nested peaks for the p100 peptide. Likewise, comparing the MS/MS of the p105 peptide in Figure 2.10 with the SRM trace in Figure 2.19, y6 > y9 ≈ y7. The relative abundance of fragments is a useful characteristic to support a signal's identity. In the absence of a heavy reference peptide co-eluting, more transitions should be included in the method to increase confidence, in essence by reproducing more datapoints of the MS/MS fragment spectrum.

In the next step various adjustments to the instrumental parameters are made, particularly the MS/MS collision energy, in order to maximize the method's sensitivity. Optimization of the declustering potential yielded a relatively flat response (Fig. 2.6) across the range, making selection of a single value for all transitions a sensible choice. Similar flat optimization of this ion source voltage has been noted for other QQQ systems such as the Thermo Quantiva [158]. An advantage of using a single value across the transition list is that the ion source remains in a constant state throughout the acquisition. Such voltage switching on the 5500 system requires 5 ms of settling time for the electronics before the next transition can be monitored. This penalty can quickly consume valuable duty cycle, the time required to cycle through all transitions in the method and record a data point for each. The duty cycle is

restricted so that sufficient data points are recorded to adequately define the chromatographic peak [116].

Repetitive acquisition of the SRM transitions while ramping the collision energy provides the best opportunity for tuning the method to increase sensitivity. Mathematical equations exist to predict collision energy, but these only take into account the peptide's precursor mass and charge state [120] [169]. The influence of amino acid composition and sequence is neglected, meaning a single collision energy value is assigned for all transitions from a peptide (gray dashed line in Figs. 2.7 and 2.11). However this is not the true case, as the different amino acid side-chains with their varying chemical properties do exert an influence on the lability of the peptide bond, as the curves in the CE ramp plots indicate that breaking the different peptide bonds requires different voltages to achieve maximum signal. The optimization of the collision energy should be considered essential in cases where sensitivity is a concern, and while a synthetic peptide to perform these scans may be considered a luxury, early effort spent in method development is rewarded downstream with improved detection. The Dynamic SILAC-SRM study strained the methodology in two ways. First, additional transitions were added to the method to monitor the new species, the M-Arg-6 labeled peptides produced from *de novo* synthesis, which reduced the dwell times for each transition by 1/3 to maintain a fixed duty cycle. Second, the earliest-possible detection of *de novo* M-Arg-6 peptides and conversely the latest-possible detection of L-Arg-0 precursors p100 and p105 in the stimulation time course demands the maximum sensitivity to confidently distinguish these weak signals from the background.

4.4 The advantages of SRM-MS over Western blotting.

In a 2013 editorial for Molecular and Cellular Proteomics [170], Aebersold *et al* compared and contrasted Western blotting to SRM-MS assays, arguing that the data derived from the MS-based targeted proteomics technique surpasses that from Western blotting on several grounds. Regarding the detection and quantification of the NF-κB precursors and their products, SRM-MS is superior to Western blotting with respect to :

i) The Quality of the Assay, including its performance characteristics, and

ii) The Quality of the Results.

4.4.1 SRM-MS vs. Western blotting: Quality of the assay.

Detection of the precursors p100 and p105 and their products p52 and p50 for the vast majority of studies has depended upon antibodies. Antibodies specific to the protein N-termini detect both the precursor/product pairs as the target epitope is shared between them, and they can be distinguished by their apparent molecular weight on the gel. C-terminal specific antibodies detect only the full-length precursor molecules. These are commercial antibodies sourced from major suppliers, but little information is provided about their properties and the consumer generally has no quality control but faith. Variations between batches of antibody as well as how it is handled and stored within the laboratory can lead to shifting performance traits, such as effective concentration and affinity [171]. Meanwhile, the synthetic peptides purchased from JPT were aliquotted in precise amounts, facilitated by the fluorescent tag at the C-terminus, and arrive with an MS spectrum and LC-UV trace to support identity and purity. An independent verification of the identity of the peptide is easily achieved on the user's end, as MS and MS/MS spectra will be ample proof of the peptide's sequence and purity. The linearity of the SRM assay is demonstrated in Figure 2.13, as the quantified synthetic peptides allow a precise titration of the instrument's detector response. While other groups commonly employ standard addition [158], supplementing different amounts of protein digest to a fixed amount of internal standard, the reverse standard addition method employed here maintains a constant background of whole cell digest, so the same amount of protein is injected onto the column each time [172]. This negates effects that increasing amounts of protein would have on column capacity and ionization efficiency. The SRM method's superior sensitivity relative to antibody-based detection is evident for the p100 peptide in the context of the p105 KO MEF stimulation time course (Western blot in Figure 1.2B vs. bar plot in Figure 2.24A). The band assigned to p100 vanishes after 4h of LTβR stimulation, while the SRM-MS method detects p100 in this assay up until the endpoint of 6h. In this case the Western blot had led to the incorrect interpretation that p100 had been largely depleted from the cell.

The Dynamic SILAC-SRM experiments are an ideal example of the additional value one obtains using MS analysis that cannot be supplied by immunoblotting. In a single measurement, two populations of proteins, the pre-existing L-Arg-0 and the *de novo* synthesized M-Arg-6 are not only distinguished but also quantified. In the multiplexed analysis, 12 peptide species (L-Arg-0, M-Arg-6, and H-Arg-10) are monitored with three transitions each, with dwell times ranging from 50 to 100 ms, a total of 36

channels measured for every 2.5 s duty cycle. It should be noted that SRM-MS permits much higher levels of multiplexing, depending on the target proteins. Monitoring a peptide from an abundant structural protein such as actin requires far less time: 2 ms is sufficient to acquire a strong signal of 1e5 cps or greater. Assuming the same duration of duty cycle, one could monitor hundreds of equivalent transitions in a single run, a level of multiplexing unrivalled by any classic immunoblotting approach.

4.4.2 SRM-MS vs. Western blotting: Quality of the results.

SRM-MS detection and quantification of proteins through tryptic peptide proxies generates datasets richer in information and more confident than those derived from Western blotting. The latter depends upon a single signal, manifested as a band appearing in a position that corresponds to the molecular weight, or more correctly the electrophoretic mobility, which is externally calibrated via a molecular weight ladder in another lane. It is not at all unusual for more than one band to appear in a lane, although this may not be shown in a published figure. At best, additional bands are an example of the antibody's non-specificity, and they can simply be disregarded based on their apparent molecular weight. At worst, these false positives cast doubt upon a band's true identity and complicate discussions about how to interpret the blot.

With SRM-MS, detection of the proteotypic peptide proxy for a protein depends on multiple parameters. One could consider the retention time of a peptide as a parallel to the electrophoretic mobility. They are of course unique: the first is a function of a peptide's affinity to non-polar stationary phase, while the second a function of the physical size of a denatured intact protein. When considered simply as fractionation methods, it is clear that the retention time is a characteristic measured at a higher resolution and therefore more precise. Whereas a typical lane in an SDS-PAGE gel might be divided into 10-20 molecular weight slices or bins, reversed-phase nanoflow chromatography might resolve hundreds of peaks over the course of a gradient elution, tuneable by manipulation of the gradient's profile and duration, making the retention time a more meaningful piece of information.

The ions isolated by Q1 and Q3 are another parameter to be weighted. Each SRM transition itself is essentially a dual mass filter, the first quadrupole isolating peptide precursors through a 0.7 m/z window, and after fragmentation of the ions transmitted, the third quadrupole allows fragments matching another 0.7 m/z wide window to

reach the detector. While this confers substantial selectivity, the complexity of a digested proteome is so extreme that despite these tight restrictions many other signals arise from other non-target peptides. Inspection of any of the chromatograms in Figures 2.14 through 2.23 illustrates the noise in these measurements, as each transition trace yields multiple peaks across the gradient. Monitoring multiple transitions per peptide and having knowledge of the relative intensity of each transition increases this selectivity further, and coincidence of all transitions when the targeted peptide elutes gives rise to the characteristic nested peaks in the correct proportions. This identification becomes indisputable when an isotopically-labeled heavy reference peptide has been spiked in, as it co-elutes with the exact retention time and peak pattern. The exquisite selectivity of the method to monitor the precursors p100 and p105 and their products p52 and p50 was demonstrated in this work by applying the SRM-MS methods to the analysis of whole cell tryptic digests of wild-type MEFs, where all proteins are present, as well as knockout cell lines where one target has been deleted ($p100^{-/-}$, $nfkb2^{-/-}$, $p105^{-/-}$, and $nfkb1^{-/-}$). The complete absence of nested peaks at the indicated retention time windows in Figures 2.16, 2.18, 2.20 and 2.22 is the supreme evidence for the method's selectivity, demonstrating that no underlying signal is contributing noise to the measurement.

The clearest advantage of SRM-MS over Western blotting is the ease with which precise quantification is facilitated when employing a heavy reference peptide. In terms of internal standards, a co-eluting molecule that is identical in every way save for the addition of a few neutrons cannot be matched, and is certainly superior to using a potentially misleading housekeeping protein such as actin or tubulin [173] [174]. The peak area of each analyte transition is normalized to the peak area of its corresponding internal standard transition. As the heavy peptides were spiked in to the protein lysate, their signal can correct for sample preparation parameters such as non-specific binding losses and digest efficiency, as well as variables at the instrument level such as injection volume and detector response. Quantitating a Western blot can be achieved [175] [176], but it requires significant work to control factors such as the linear detection range and sample loading, and relies on densitometry, the counting of pixels, that is itself highly variable [177]. However studies have shown that it is the operator itself that is the biggest source of error due to differences in sample handling [178].

4.5 Dynamic SILAC SRM-MS analysis of LT-βR stimulated MEF cells.

The successful quantitation of the NF-κB precursors and products in MEF wild-type and knockout cell lines was a proof-of-concept to show that SRM-MS analysis can achieve what Western blot analysis had done previously. This is best demonstrated by the 12 h stimulation time courses in Figure 2.25. In whole cell extracts from wild-type MEF cells, a gradual, stimulation-dependent decrease in the levels of p100 and p105 are clear, adhering to the established hallmarks of non-canonical signaling. This is matched by the slow rise of the products p52 and p50. For each species every timepoint is comprised of six data points representing two transitions quantified for three separate injections, the clustering of the data points a measure of the precision of the measurement. Precisely quantified signals permit the calculation of fold-changes for a single species between any time points, and importantly the absolute nature of the quantification means the levels of the different species can be compared directly. Absolute quantification transforms the qualitative thinking about the system, that a band in lane 1 on an immunoblot is stronger or weaker than that in lane 2, into a capacity to think about the very real molecules that are the actual players in this cellular signaling mechanism. With absolute quantification a 3-fold excess of p100 over p105 becomes apparent, and one can observe a loss of 200 million molecules of p100 (per µg protein) between 0 and 12 h, with a rise in its N-terminal product p52 by an approximately equal magnitude. In the p100 KO and nfkb2$^{-/-}$ MEF cell lines (Figure 2.24B), SRM-MS measurements of the basal and stimulated levels of p105 and p50 reveals significant differences that are not obvious from Western blotting, ultimately leading to a better-informed mechanistic understanding of the cell lines.

The expanded capabilities of the Dynamic SILAC SRM-MS experiment is the most convincing demonstration of the value of this approach, exploiting the power of isotopic labeling for the dual purpose of absolute quantification and to mark protein populations relative to the initiation of non-canonical activation. This strategy further transformed our interpretation of the mechanism of p100/p105 activation and presented a new viewpoint that would have eluded our detection otherwise. Preferential degradation of the old L-Arg-0 precursors combined with preferential processing of *de novo* M-Arg-6 precursors can account for the slow kinetics of the activation, as it is only after new precursor has been translated at the ribosome and processed by the proteasome that the rise in products occurs and the transcriptional

response to the stimulus enacted. This is also consistent with a co-translational mechanism that was proposed within the Scheidereit group 14 years ago by Mordmueller *et al*. [179], as the kinetics would suggest that the nascent precursor polypeptide chain emerging from the ribosome is immediately acted upon by the proteasome, p97 and unknown cofactors to generate the product. The requirement for protein synthesis to upregulate NIK and the essential role of the proteasome have been established [16] [15] so it was unsurprising that inhibition with cycloheximide or MG-132, respectively, should prevent the accumulation of the products p52 and p50, but through Dynamic SILAC it became clear that it was specifically the *de novo* M-Arg-6 products that are affected. The investigation into p97's role in this process was the most illuminating and opens up new avenues for ongoing investigations. Blocking the p97 ATPase activity by inhibition with DBeQ significantly reduced the generation of *de novo* products, leading to a speculative mechanism for a long-standing question in the NF-κB world: the mechanical force created by p97 hydrolysis of ATP is harnessed and used to pull the N-terminal halves of the precursors p100 and p105 free of the proteasome before they are completely degraded; this force may also be required in separating the tightly-bound NF-κB dimers

4.6 Modelling predicts a signal-responsive precursor complex.

The interaction and interdependence of the precursors p100 and p105 in non-canonical signal propagation pointed to the possibility of a functional p100-p105 complex existing. The development of a mathematical model to describe this aspect of non-canonical signaling, the conversion of the free or complexed precursors p100 and p105 into their respective products, benefitted greatly from the absolute quantitation of protein species through a 12 h time course. Previous models published on NF-κB have focused on canonical signaling alone [180] [181], with modelers compiling data and deriving rate constants for reactions in the pathway from across the literature. This means that ultimately the data is sourced from different researchers in different groups working in different labs with different cell lines at different times and different places. In our work, all of the data inputted into the model was derived from our own measurements, performed with careful coordination and consistency. As has been discussed, the absolute data from SRM-MS analysis is of a higher quality, and measurements of whole cell lysates in endogenous systems brings us closer to the physiological truth. The proposed models in Figure 2.31 were

kept simple, focused on this single aspect of precursor conversion and omitting upstream processes like receptor activation and downstream processes such as nuclear translocation of the products. SRM-MS measurement of cells with and without stimulation provided the concentration changes that occur over time, so kinetic parameters could be estimated by fitting each model to this absolute, time-resolved data. The mathematical modeling efforts supports the concept that p100 and p105 form a complex and that this complex responds to the non-canonical stimulation. Model M3 emerged as the most probable model, in which a p100-p105 complex forms and responds to the non-canonical stimulus by undergoing proteasomal processing. Simulations performed with this model were predictive and agreed well with the experimental data as evident in Figure 2.33, and the model also predicted that 14% of total p100 and 42% of total p105 participate in a p100-p105 complex. Matthias Wilm once wrote [182]: “A model remains a model, it never describes the reality in an absolute way. A good model should explain experimental data and should have a good predictive power for experimental results in a qualitative or even quantitative way. A model is not true because it is believed to be true. It is even difficult to say that certain experimental findings confirm the model description. It is more accurate to say that the experimental data can be interpreted within the framework of a specific model.” While Wilm was referring to models of electrospray ionization, this concept is equally apt here. While the model alone cannot be taken as proof of a functional p100-p105 complex, it is a useful step, and provides a framework for evaluating future findings. Much like models of the canonical system [181], this model of non-canonical NF-kB activation will be subject to ongoing evolution and challenge as future experiments attempt to further dissect an elaborate mechanism and generate data that may support or refute the model’s fundamental construction.

The success of these efforts encourage the continued development of mathematical models to explain aspects of NF-κB signaling involving more of the upstream and downstream proteins. To this end SRM methods have been created to cover other transcription factors (Rel A, Rel B, c-Rel) as well as the IKK kinase complex. It would be interesting to incorporate Dynamic SILAC SRM-MS data of stimulation time courses in other cell types or with different stimuli. Ultimately, a larger and more comprehensive model of the non-canonical NF-κB pathway that gives deeper mechanistic insights into the kinetics could find utility in the development of therapies targeting this system, pinpointing reactions that might be pharmacologically

manipulated to achieve a desired correction to pathway misregulation occurring in disease states [183] [184] [185].

4.7 Protein-protein interaction analysis of p100 and IKKY/NEMO.

Analysis of a LTBR stimulation time course in MEF cells by Dynamic SILAC SRM presents an interesting question: if the two populations of precursors, old and new, exhibit distinct behaviours, how does the cellular machinery distinguish one from the other? One possibility is that in each context the precursors associate with different proteins, and this protein interaction network guides them down the appropriate pathway. An MS-based proteomics approach to find significant interactors of p100 was not able to shed much light on this question. Either SILAC or LFQ was used as the basis for normalization to quantify differences between the immunoprecipitations with an anti-p100 anitbody versus a control IgG. Both approaches pointed to the other NF-κB molecules as the primary significant interactors of p100: p105 (nfkb1), Rel A (nfkb3), Rel b and, to a lesser extent, c-Rel. That p105 was identified as a significant interactor is expected given their previous co-occurrence in Co-IP experiments, the known interdependency of processing of the two precursors, and our mathematical model's assertion that they form a functional complex. It is known that the C-terminal ankyrin repeat domain of p100 plays an inhibitory role by binding to the NF-kB transcription factors and masking their nuclear localization signals to maintain them in the cytoplasm in an inactive state [14]. It is clear from these data that no other protein binds p100 with as much affinity. The AAATPase p97 was equally abundant in both the experimental and control pull-downs, therefore it cannot be classified as a significant interactor of p100 according to this approach. Considering that p97 has been demonstrated here as having a functional connection to p100 biology, regulating the proteasomal processing of the two precursors, this might seem surprising. However, a failure to enrich p97 in the p100 pull-down relative to IgG could be explained by the difference in abundance. While p97 is highly abundant, by some measures accounting for 1% of total protein mass [186], p100 is relatively scarce. Even if p100 bound to p97 with high affinity and equal stoichiometry, this would not decrease the pool of unbound p97 in a significant way. It is likely that p97 mediates its interaction with p100 through an as-yet-unknown cofactor, although the MS screen did not enrich for any likely candidates.

The MS-based screening for interactors of IKKγ/NEMO successfully identified a p-body protein, enhancer of decapping 4 (EDC4), which was subsequently shown to undergo an IKK-mediated phosphorylation in response to DNA damage caused by radiation, and has a function in regulating mRNA transcript stability (Mikuda *et al*, submitted). This exciting finding points to new roles for the IKK complex in cellular regulation of mRNA that extend beyond its classical roles in NF-κB signaling, with RNAseq data pointing to hundreds of affected transcripts, the large majority of which are not known NF-κB target genes. The identification of EDC4 would not have been possible without an MS analysis of the eluate from the co-immunoprecipitation of NEMO. The MS-based detection of proteins pulled down by affinity enrichment is completely unbiased; the bound proteins are eluted, digested and their peptides measured by a data-dependent shotgun analysis. Detection by Western blot, conversely, is necessarily biased, as specific antibodies are required to probe the blot for proteins suspected of having an interaction with the bait. Cost and practical limits on how many times a blot can be probed restrict these to a handful of interactors. This limitation also shields an ugly truth of Co-IP experiments, that the total number of proteins pulled down routinely exceeds 1000 and often many more. In the p100 examples here, nearly 3000 proteins occupy the central cluster of non-enriched background proteins, while for the NEMO screen this number exceeded 1800. Optimization of experimental conditions, particularly the composition of the post-IP washing buffer, can help to reduce this number but such efforts are often offset by fear of washing away something important. The high number of binders pulled down is not surprising considering how the protein input for the IP is prepared. For both the p100 and NEMO IPs, cells are harvested from culture and lysed by shearing outer membranes in a douncer, after which the cytoplasmic fraction was isolated. At this stage the cellular context in which the proteins existed has been destroyed: the contents of various compartments intermingle, protein structures and scaffolding have collapsed, and the remaining protein soup is assayed to try and fish out the bait protein and whatever proteins are bound to it, or bound to the anti-body, or to the beads to which the antibody is linked. Proteins that might never share a cellular locale can co-precipitate, resulting in the aforementioned massive background and possibly confounding researchers with false positive hits. This protein lysate is far removed from the original cellular conditions, in which the protein concentration is estimated to be between 100-300 mg/mL [124], an astonishing figure achieved by the tight packing and compartmentalization of the cell.

Alternative approaches allow researchers to get closer to a protein's interactors under *in vivo* conditions. Labeling proteins with fluorescent groups and performing microscopic imaging to demonstrate co-localization [187], this method suffers from a similar bias, in that only those proteins so tagged can be studied. Such fluorescent proteins are often overexpressed and that, in combination with the tagging, means the system can no longer be classified as endogenous. Furthermore, the low resolution attained by imaging is not definitive proof of an interaction. Both of these issues are addressed with the proximity ligation assay [188] [189], a technique to image protein-protein interactions in a cell, whereby antibodies to two proteins are tagged with short DNA strands. If the two proteins come into contact, when they are recognized by their respective antibodies the two DNA strands can participate in a rolling circle synthesis, generating circular DNAs amplified hundreds of times. Detection occurs through hybridization with fluorescently labeled oligonucleotide probes, yielding a signal visible to the microscope. While addressing proximity at endogenous levels, the requirement for antibodies to proteins being investigated makes this approach biased.

Protein cross-linking to preserve complexes or investigate protein structure has been present in the literature for some years, but needed refinement on the technical and bioinformatics fronts to make it feasible [190] [191] [192]. It is an MS-based approach whereby cells in culture are treated with a small-molecule chemical reagent that covalently links proteins together. Typically these are based on bifunctional molecules such as two aldehydes or two NHS esters connected by a short hydrocarbon linker [193] [194]. The functional groups form a bond to protein primary amines (N-terminus, Lys) that have the correct proximity to each other, through a reductive methylation or simple substitution reaction. After cross-linking, the cells are lysed and the target enriched through affinity purification, after which the proteins are prepared for a standard shotgun analysis by tryptic digestion. Where crosslinks between two proteins formed, digestion yields two tryptic peptides connected by the short hydrocarbon linker. MS/MS fragmentation of such species gives complex patterns, but computational tools have been developed to aid in their interpretation [195]. Proteins cross-linked together is evidence of a primary interaction, and one can map the peptides onto known structures to determine the domains making contact. A core of the complex can be described and re-asssembled outwards as secondary and tertiary cross-links are defined. Unlike the MS-based approaches utilized in the present work, significantly enriched interactors can be further ranked according to

their proximity to the target. This methodology has been used to characterize protein-protein interactions within the 50 mDa yeast nuclear pore complex [193], sorting interactions as local, distant and transient. A similar approach to elucidation of the interactome of p100 might meet with more success, capturing binding weak or transient interactors at different points during a LTBR stimulation time course. As recent reports indicate that the majority of cellular protein networks are governed by weak interactions with very few stable core complexes [196], cross-linking may be the only way to preserve these interactions for study.

4.8 Polyubiquitination in cell signalling networks.

The modification of proteins with ubiquitin polymers is a critical PTM, controlling the fate of marked proteins and acting as a major regulator within signaling networks [101], with NF-κB the archetype. In fact, it was the first system in which the degradative role of K48-linked polyubiquitination was described, activating the canonical pathway by down-regulating the inhibitory IkBα, or suppressing NIK and activating p100 processing in the non-canonical pathway [197] [198] [199]. While K48-linked polymers are the classical example, other major polyubiquitin types exist within this system. Non-degradative K63 chains are needed for canonical NF-κB response to the chemokine IL-1B [200], while K63 chains form on TRAF6 following genotoxic stress and are essential to the DNA damage transcriptional response by promoting IKKβ phosphorylation [126]. K11 chains can act as degradative signals, as with the anaphase promoting complex [201] or, conjugated to RIP1, are also involved in NF-κB activation [202]. As for linear (M1) polyubiquitin, one of the only substrates shown has been IKKY/NEMO, where it is involved in the response to TNFα [133]. Identification of chains via Western blot is possible, but the antibodies available are generally of low affinity and specificity [203]. Methods based on SRM-MS detection have been developed that can overcome these limitations using isotope-coded reference peptides that both increase confidence and aid in quantification [128]. One drawback of such a proteomic approach is the loss of context that occurs upon tryptic digestion, separating the chain from the substrate and thereby obscuring its identity. Here a combination of Western blotting and MS-analysis can be useful to pinpoint the primary substrate.

4.8.1 Linear ubiquitination and parkin.

SRM-MS analysis of ubiquitin linkages enriched from HEK293T cell lysate through affinity to a recombinant NEMO UBAN domain was used as evidence for parkin's role in M1 chain formation [130]. Cells over-expressing wild-type parkin in combination with the LUBAC components HOIP and HOIL had 3 times as much linear chain formation, an effect abolished by deleting parkin's UBL domain. Importantly, knockdown of parkin by RNA interference could demonstrate that this was also the case at endogenous levels. Parkin's involvement in linear ubiquitination was further cemented by it displaying higher affinity for HOIP than a known component of LUBAC, Sharpin [133], via an MS-based interactomics experiment that compared proteins enriched through co-immunoprecipitation with HOIP or HOIL. Overall, this work linked NF-κB signaling, activated by linear ubiquitination of NEMO, to a stress-protective pathway and mitochondrial integrity through up-regulation of OPA1, an NF-κB target gene.

4.8.2 Novel regulation of protein kinase A signalling.

Mass spectrometry was key in elucidating two new regulatory mechanisms inhibiting signaling through Protein Kinase A (PKA), an important signaling hub whose misregulation has implications in disease [187]. A previous proteomics experiment identifying binding partners of the catalytic subunit of PKA, PKAC, motivated investigations into what role one of these binders, ARHGAP36, plays in PKA signaling. Imaging studies indicated that the presence of ARHGAP36 led to lysosomal degradation of PKAC. An SRM-MS analysis of a PKAC pull-down showed enhanced formation of K63-linked chains in the presence of ARHGAP36, with these chains later proven essential to targeting PKAC to the lysosome for destruction. A second mode of inhibition is through ARHGAP36 acting as a pseudosubstrate to bind and block the catalytic site of PKAC. Here a quantitative shotgun MS experiment based on IBAQ [139] could show that the two proteins exist in the same abundance range. This is an important finding as an equivalent stoichiometry is a prerequisite for pseudosubstrate inhibition to be a feasible mechanism.

4.9 Quantitative proteomics and unbiased biology.

Quantitative MS-based proteomics has revolutionized protein biology over the last three decades. The scale on which proteins are reliably detected and confidently identified has greatly expanded the protein universe, facilitated by the arrival of complete genome databases [66]. Quantitation of proteins is an important way to study a protein's function, its protein interactors, and how perturbation of the system can alter its expression [111]. Confined to classical molecular biology techniques with their reliance on antibody-based detection, such studies were necessarily confined to a limited scope. Such biased approaches often meant that one did not detect a protein unless they are actively looking for it. Modern MS-based proteomic studies analyzed by a data-dependent acquisition scheme are unbiased and capable of monitoring many thousands of protein species in a single measurement.

Quantitation in MS-based proteomics, including the majority of studies presented in this thesis, utilize an isotope-coding strategy to synthesize peptide analogues that are identical to the target peptide in every way except for their mass. As an internal standard this strategy is unparalleled, as normalizing to this signal can correct for a number of parameters covering sample preparation as well as the measurement itself [204]. When the goal is relative quantitation, the internal standard need only act as a fixed signal, a constant magnitude that does not waver across multiple samples and, ideally, has an MS intensity similar to the target. This is the most common form of quantitation and was employed in the present work in the search for protein interactors and for profiling polyubiquitin linkage types. Absolute quantitation, on the other hand, interprets the fixed signal as representing a defined amount and therefore takes the normalization further and uses it to assign concentration units to the measurement. Absolute quantitation is more powerful than relative quantitation. By assigning units, one can directly compare the levels of different proteins across samples. In the mathematical modeling of the precursors-to-products mechanism in non-canonical NF-κB signaling, the absolute values of the precursors and products were used to estimate the kinetic parameters governing the various reactions: synthesis, degradation, complex formation, and processing [25]. However, it is important to be aware that absolute quantitation by MS-based proteomics has its limitations, ranging from specific technical issues to the broader picture of cell biology.

4.10 Caveats of absolute quantitation.

Absolute quantitation of NF-κB proteins p100, p105, p52 and p50 was achieved using isotopically-heavy peptides purchased from a vendor [205]. A proprietary fluorescent C-terminal tag facilitated accurate aliquotting of peptides, delivered as 1 nanomole lyophilized in plastic tubes. The user re-solubilizes the peptides and prepares a stock solution according to a vendor-supplied protocol. During sample preparation these reference peptides are added to a solution of whole cell lysate, prepared by a Wessel-Fluegge methanol-chloroform extraction [122], prior to reduction and alkylation of protein sulfhydryl groups and tryptic digestion. At this point in the process several assumtions have already been made that impact absolute quantitation. The first is that the actual amount of lyophilized peptide in the tube corresponds with the amount claimed, and the second is that the resolubilization is 100% efficient. While the first assumption is reasonable enough to make, in my experience the process of resolubilization should not be taken for granted. Method development work on the peptides for the expanded NF-κB SRM study encountered significant problems with this seemingly straightforward step. Despite following dissolution protocols, unexpectedly weak peptide signals during MS characterization prompted the discovery that these measures were insufficient. Increasing the shaking times as well as increasing the organic solvent content improved the yield, but it is challenging to assess the extent of losses. Loss of internal standard peptide, either due to vendor error or solubilization problems, will lead to a systematic over-estimation of the target peptide in the sample. The perfect internal standard would be an entire protein labeled with heavy isotopes, quantified accurately, and existing in a stable solution [206]. Such a protein could be introduced earlier in the sample processing, alongside the initial cell pellet, and therefore correct for all handling steps including fractionation, boosting sensitivity for low-abundance molecules.

On the sample side, the protein solution the peptides are added to has already undergone several manipulations, including ultrasonication, vortexing, precipitation of proteins and resolubilization. Cumulatively, these steps alter the proteome through heterogeneous extraction, with loss of membrane proteins the most severe [207]. Such losses cannot be reasonably assessed and are therefore neglected from consideration. One also makes an assumption that the enzymatic digestion is 100% efficient, that the target peptides are completely liberated from their parent polypeptide. While Schevchenko *et al* demonstrated efficiencies of ~90% [121], each peptide exists in a unique context determined by the surrounding protein sequence.

Secondary protein structures as well as unfavourable residues adjacent to cleavage sites affect digest efficiency, although issues with the latter can be avoided with careful peptide selection during method development. One can also avoid peptides with known PTMs, but there remains a distinct possibility of an undiscovered PTM altering the peptide mass and making it invisible to the targeted MS method. From the moment the target peptide is released into solution, its isotopically-coded analogue acts as a perfect internal standard. Non-specific losses, sample loading, chromatography and instrument response will affect both peptides equally [204]. However, the number of assumptions that have been made up until this point will influence any absolute quantitation efforts. Incomplete protein extraction and digestion will lead to an underestimation of the target peptide concentration. Compounding the technical difficulties of absolute quantitation is the influence of the scientist doing the work. In one study, efforts to quantify the same proteins from identical sample solutions using identical protocols saw a 25% variation across different research institutes [208].

It is clear that the very nature of biology itself introduces significant variability into an experimental outcome. The behaviour of cells and the expression of proteins depend on their environment, and this can change tremendously from lab to lab. Cell lines with the same nominal identity may have undergone genetic divergence [209] or have been contaminated by other cells or viruses [210]. In the present work, the same MEF cells grown in normal growth media versus those grown in a depleted SILAC media responded differently to the LTβR stimulation. In terms of p100 levels, the former reached an endpoint about 3 times lower than the basal levels, while the SILAC cells underwent a 2-fold decrease in p100 levels.

With so much variability derived from the technical and biological aspects of this work, how can one know what the best strategy is? The concept of “absolute” quantitation is suggestive of some kind of universal truth that will be obtained, that Protein X will be found in Cell Y at a concentration of Z. It is impossible to know for certain. The best that can be hoped for in absolute quantitation is that the data itself will guide the way to the truth. If multiple measurements by multiple labs using multiple techniques arrive at a similar number, then eventually this will be accepted as being accurate. However one can easily imagine the alternative case, and the resulting disputes in the literature justifying one number over another. Given the broad dynamic range of protein expression in cells, perhaps absolute quantitation

should be approached with broader strokes. Confining a protein's absolute concentration to an order of magnitude can already be quite informative. The IBAQ technique used to determine the approximate stoichiometry of PKAC and ARHGAP36 in HEK293T cells is a good example of this, distributing proteins across the proteome into abundance bins. An advantage of this shotgun approach is that it is unbiased and generates a deeper dataset that could be revisited in the future. However, a targeted approach such as SRM-MS will provide a more robust measurement, and more data points acquired better outline the chromatographic peak. Such an approach is necessary for quantifying low-level protein species reliably across many samples, such as time courses or clinical samples from large-scale studies. The ease and robustness of this form of quantitation offsets the expense incurred for synthetic reference peptides. An SRM-like technique used on quadrupole-Orpitrap instruments may come to dominate in the future [211] [212]. It follows the same principles but the Orbitrap measurement offers higher resolution, mass accuracy and the simultaneous measurement of all peptide fragments. While triple-quadrupole instruments are more affordable, the quadrupole Orbitrap has already found a home in many labs specialized in shotgun proteomics.

4.11 Concluding remarks.

If the past two or three decades are any indication, the future is sure to bring new mass spectrometers with faster scanning that offer more sensitivity, mass accuracy and resolution. Proteomes will be characterized more completely and permit a deeper mechanistic understanding of protein function, particularly in critical cell-signaling pathways such as NF-κB that attract considerable interest for links to disease and therefore their therapeutic potential. Proteomic approaches will continue to be adopted on a larger scale in labs worldwide, though the expense and expertise required will still limit it to more specialized groups. These factors have always meant the proteomic community is relatively young and small, and this has fostered a certain amount of consistency to how this work is executed. What started as a Wild West of experimental strategies and bioinformatics approaches was eventually tamed and standardized, aided by the collaboration of several top labs in producing guideline publications, first for shotgun [213] [214] and later for targeted proteomics [204], of what constituted sound practices and acceptable data. A perusal of the literature shows that proteomics labs around the world now prepare routine samples using the

same protocols and measure them with similar chromatographic and instrumental settings. There is no question that this has met with tremendous success but it also represents a stagnancy that is out of step with the innovative spirit that has driven proteomics from the beginning. Indeed, there are groups disrupting these norms by exploring things like alternative enzymes to trypsin [215], MS-friendly detergents [216], or alternative chromatographic stationary phases and flow-rates. Uncovering the remainder of the proteome may require a re-examination of some of these fundamental practices that may systematically mask protein populations from detection.

5 Summary

Activation of the NF-κB precursor protein p100 and p105 by a specific proteasomal truncation to yield the active products p52 and p50 is a distinct feature of the non-canonical pathway but the mechanism governing it remains elusive. A novel mass spectrometry-based proteomics strategy was developed, using the targeted selected reaction monitoring technique in conjunction with stable isotope labeling for both absolute quantitation of proteins and to mark precursor protein populations relative to the application of the lymphotoxin β stimulation. In an endogenous murine embryonic fibroblast system, we have shown that both precursors are processed to the respective products in a parallel and interdependent manner in response to a lymphotoxin β receptor agonist. Our Dynamic SRM-SILAC method allowed distinction of pre-stimulation protein populations from proteins synthesized *de novo* post-stimulation, and revealed a tendency for older precursor molecules to undergo degradation while the de novo molecules went on to be processed to the products, accounting for the slow and persistent kinetics that are a hallmark of the non-canonical NF-κB pathway. In addition, the hydrolytic activity of the AAA ATPase VCP/p97 was implicated in the generation of de novo p52 and p50.

An MS-based proteomics screen for specific, radiation-induced protein interactors of another key NF-κB player, the regulatory subunit of the IKK complex, IKKγ/NEMO, turned up the Enhancer of mRNA Decapping 4 (EDC4). This unexpected finding has expanded the known role of NF-κB regulation of protein levels beyond transcription into mRNA stability. Separate investigations into the ubiquitin E3 ligase, parkin, connected it to the NF-κB pathway through a linear ubiquitination it helps catalyze on IKKγ/NEMO. Proteomic analysis of polyubiquitin in the protein kinase A (PKA) signalling pathway helped to identify two novel modes of regulation: a lysosomal degradation pathway; as well as a pseudosubstrate inhibition mechanism.

Keywords:

Targeted proteomics, NF-κB signalling, selected-reaction monitoring mass spectrometry, mathematical modeling, polyubiquitination.

6 Zusammenfassung

Die Aktivierung des NF-kB Vorläuferproteins p100 und p105 erfolgt durch eine proteasomale Trunkierung, um die aktiven p52 und p50 zu erzeugen. Zur Erforschung wurde eine Massenspektrometrie-basierende Proteomik-Strategie entwickelt, die mit der gezielten Reaktionsüberwachungstechnik plus einer Isotopenmarkierung verwendet wurde. In einem endogenen murinen embryonalen Fibroblasten-System konnte gezeigt werden, dass beide Vorläufer zu den jeweiligen Produkten in einer parallelen und sich einander bedingenden Weise in Reaktion auf einen Lymphotoxin-β-Rezeptor-Agonisten verarbeitet werden. Unsere SRM-SILAC-Methode erlaubte die Unterscheidung von Prä-Stimulationsprotein-Populationen aus Proteinen, die de novo nach der Stimulation synthetisiert wurden, und zeigte eine Tendenz für ältere Vorläufermoleküle, sich einer Degradation zu unterziehen, während die de novo-Moleküle auf die Produkte verarbeitet wurden. Die langsame und anhaltende Kinetik, die ein typisches Merkmal des nicht-kanonischen NF-κB-Weges ist. Darüber hinaus, konnten wir beobachten, dass die hydrolytische Aktivität der AAA ATPase VCP / p97 in der Bildung von de novo p52 und p50 eine Rolle spielt. Durch ein MS-basiertes Screening für strahlungsinduzierte Protein-Interaktoren eines weiteren NF-κB-Players, der die regulatorische Untereinheit des IKK-Komplexes IKKγ / NEMO bildet, konnte der Enhancer von mRNA Decapping 4 (EDC4) entdeckt werden. Durch die zusätzliche Untersuchung der E3-Ligase, Parkin, konnte eine Verbindung mit dem NF-κB-Weg durch eine lineare Ubiquitinierung hergestellt werden. Es konnte gezeigt werden, dass Parkin ist Hauptkomponenten des linearen Ubiquitin-Ketten-Assemblierungskomplexes (LUBAC). Die Proteomanalyse im Proteinkinase A (PKA) -Signalisierungsweg konnte zwei neuartige Regulationsformen identifizieren: K63-verknüpfte Polyubiquitinierung die katalytische Untereinheit von PKA, PKAC in Richtung eines lysosomalen pathways führt, und auch durch einen Pseudosubstrat-Hemmungsmechanismus.

Stichworte:

Gezielte Proteomik, NF-κB-Signalisierung, selektierte Reaktionsüberwachung Massenspektrometrie, mathematische Modellierung, Polyubiquitinierung.

7 Materials and methods.

7.1 Materials.

7.1.1 Chemicals.

Acetic anhydride (Sigma-Aldrich, Germany)

Acetonitrile, Chromasolv LC-MS grade (Fluka, Germany)

Ammonium bicarbonate (Roth, Germany)

Ammonium persulfate (Sigma Aldrich, Germany)

Arginine $^{13}C_6$ & $^{13}C_6{}^{15}N_4$ (Cambridge Isotope Laboratories, USA)

Benzotriazol-1-yl-oxytripyrrolidinophosphonium hexafluorophosphate, PyBOP (Merck, Germany).

T-butyl methyl ether (Sigma Aldrich, Germany)

2-chloroacetamide (Merck, Germany)

Dimethylformamide (Sigma Aldrich, Germany)

Dithiothreitol (Sigma Aldrich, Germany)

Empore C18 solid-phase extraction disc (3M, Germany)

Formic acid (Sigma Aldrich, Germany)

Fmoc amino acid derivatives (Intavis, Germany)

4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid, HEPES (Sigma Aldrich, Germany)

Isopropanol, Lichrosolv (Merck, Germany)

L-Lysine-D4 (Cambridge Isotope Laboratories, USA)

L-Lysine $^{13}C_6{}^{15}N_2$ (Cambridge Isotope Laboratories, USA)

Methanol, LC-MS grade (Merck, Germany)

N-methylmorpholine (Intavis, Germany)

N-methyl-2-pyrrolidone (Sigma Aldrich, Germany)

Piperidine (Sigma Aldrich, Germany)

Reprosil-Pur AQ C18 chromatography beads (Dr. Maisch, Germany)

Rotiphorese acrylamide solution, 30% (Roth, Germany)

Sodium dodecylsulfate (Merck, Germany)

SpikeTides TQL synthetic peptides (JPT Peptide Technologies, Germany)

Tetramethylethylenediamine, TEMED (Sigma Aldrich, Germany)

Thiourea (Roth, Germany)

Tris(2-carboxyethyl)phosphine, TCEP (Sigma Aldrich, Germany)

Trifluoroacetic acid (Sigma Aldrich, Germany)

Urea (VWR, Belgium)
Water, LC-MS grade (Merck, Germany)

7.1.2 Enzymes.

Benzonuclease (Pierce, USA)
Endoproteinase Lys-C (Wako, Japan)
Sequencing-grade trypsin (Promega, USA)

7.1.3 Laboratory equipment

Analytical balance XA105DU (Mettler Toledo, Germany)
Benchtop Centrifuges 5415D and 5430R (Eppendorf, Germany)
Capillary tubing laser puller (Sutter Instruments, USA)
Cold trap solvent recovery system (Thermo Savant, USA)
Fused-silica capillary tubing (Polymicro technologies, USA)
Hamilton syringe needle, point style 3 & plunger (Hamilton, Switzerland)
Heating block DB2A (Techne, UK)
Microscope Sedival (Carl Zeiss, Germany)
Mini-PROTEAN electrophoresis system (Bio-Rad, Germany)
Pipet-lite XLS pipets: P2, P10, P20, P200, P1000 (Rainin, USA)
Pressure bomb for capillary column packing (Biostep, Germany)
Sonifier 450 ultrasonic cell homogenizer (Branson, USA)
Thermomixer incubator/shaker (Eppendorf, Germany)
Vortex Genie 2 (Scientific Industries, USA)

7.1.4 Instrumentation.

Nano-LC-ultra 1D plus (Eksigent, USA)
PAL HTC-xt autosampler (CTC, Switzerland)
Proxeon Easy nanoLC1000 (Thermo Scientific, Denmak)
Q-Trap 5500 triple quadrupole MS (Sciex, Canada)
Q-Trap 6500 triple quadrupole MS (Sciex, Canda)
Q-Exactive (Thermo Scientific, Denmark)
ResPep SL solid-phase peptide synthesizer (Intavis, Germany)

7.2 Methods

7.2.1 Solid-phase peptide synthesis of candidates for NF-κB assay.

Synthetic peptides were prepared to aid in the development of selected-reaction monitoring (SRM) MS methods using a solid-phase synthesis approach on an automated peptide synthesizer, the ResPep SL (Intavis). The chosen solid-support, Tentagel Amide resin, has a modified polystyrene core and is used to generate peptide carboxamides. At a scale of 5 µmols synthesis target, a slurry of 22 mg of the Tentagel Amide resin (capacity 0.23 mmol/g) in 200 µL dimethylformamide (DMF) was required for each peptide. Peptides are synthesized from the C-terminus towards the N-terminus in a cycle involving activation of the carboxylic acid with benzotriazol-1-yl-oxytripyrrolidinophosphonium hexafluorophosphate (PyBOP) followed by deprotection with piperidine to prepare the added residue for the next amino acid in the sequence. Acetic anhydride is introduced to scavenge and react with and neutralize any unreacted amine groups, and extensive washing with DMF is performed between each cycle step to eliminate carryover of reagents and amino acid derivatives. The amino acids used are Fmoc-protected derivatives, a chemistry used to block the reactivity of amino acid amine groups. The derivatives were all prepared as 0.5M stocks in DMF, with the exception of His, Phe and Pro which are dissolved with the aid of a sonication bath in N-methyl-2-pyrrolidone. The carboxylic acid activator, PyBOP was dissolved in DMF and a fresh solution prepared after 30 h to compensate for degradation effects. Following coupling, deprotection of backbone amines is achieved with piperidine. Upon completion of synthesis, the resins with peptides bound were washed extensively with ethanol and dried, followed by treatment with a solution of 92.5% TFA and 5% triisopropylsilane that cleaves the peptide from the solid-phase support to yield the peptide amide, while simultaneously deptrotecting reactive side-chains blocked with t-butyl and Boc groups. Following cleavage from the resin, peptides were precipitated from solution with cold t-butyl-methyl ether and the white solids washed with multiple aliquots of more ether.

7.2.2 Quantified SpikeTides TQL peptide standards.

Peptides were delivered as 1 nmol lyophilized aliquots, the quantitation achieved through a proprietary C-terminal fluorescent tag. All peptides are tryptic and have a C-terminal lysine or arginine residue isotopically coded as either Lysine $^{13}C_6{}^{15}N_2$ (Lys8) or Arginine $^{13}C_6{}^{15}N_4$ (Arg10) and a typical proteomic workflow with trypsin digestion will cleave off the fluorescent tag. Peptides were dissolved at a concentration of 10 pmols/µL in 20% acetonitrile (ACN) with 100 mM ammonium bicarbonate buffer (ABC, pH 7). Longer peptides with more hydrophobic character were dissolved in 50% ACN/100 mM ABC. Vortexing and sonication in a water bath were employed to promote complete dissolution and peptide stocks stored at -80 °C.

7.2.2.1 SpikeTides in an in-solution digest.

The required amount of SpikeTide was added to the protein lysate at the beginning of the workflow. For the absolute quantitation of the precursors and products in the NF-κB pathway, a working solution of the heavy peptides at 25 fmols/µL was prepared and 4.8 µL of this was spiked into 120 ug of MEF cell lysate.

7.2.2.2 SpikeTides in in-gel digest workflow.

The SpikeTides are introduced at the end following peptide desalting, as they are too small to be retained by the polyacrylamide matrix. In this context the SpikeTides are used for relative quantitation and must be pre-digested with trypsin to remove the C-terminal tag. A portion of the 10 pmol/µL stock solution was diluted with 50 mM ABC buffer to a final ACN content of 7%, which is within the tolerance of trypsin. Add 1 µg of trypsin and incubate with gentle shaking for 4 h at room temperature. Stop digest by adding TFA to a final concentration of 1%, then capture and desalt peptides using a StageTip. Wash peptides and elute with 80% ACN, lyophilize and reconstitute peptides in a buffer containing 3% ACN and 0.1% formic acid at a concentration of 2 pmols/µL. To use for relative quantitation, reconstitute samples with buffer containing 50-100 fmols/µL of heavy SpikeTides.

7.2.3 StageTips for peptide enrichment and desalting.

Stop-and-Go Extraction Tips are homemade soild-phase extraction devices prepared according to Rappsilber *et al* [217]. Punch three discs from 3M Empre C18 material with a Hamilton blunt-tipped syringe needle and deposit discs in a P200 pipet tip, using the syringe plunger to tamp them down and secure a leak-proof fit. Liquids are introduced through the top and the tip is placed through a hole punched in the lid of an Eppendorf tube acting as a receiver. Centrifugal force generated by spinning for 2 min at 5000 rpm in a benchtop centrifuge pushes liquid through the C18 material and it collects as flow-through in the receiving tube. Prior to use, the tip is wetted with 50 µL methanol and equilibrated with 100 µl 2% ACN/1% TFA. Samples (e.g. cell lysate digest) are acidified to 1% TFA and spun at maximum speed in a benchtop centrifuge to pellet insoluble material. After sample loading, the captured peptides are washed with 2% ACN/1% TFA prior to elution with 80% ACN/0.1 % formic acid. Following lyophilization in a SpeedVac, samples are reconstituted with 3%ACN/0.1% FA.

7.2.4 Preparation of cell pellet for enzymatic digest.

Cells derived from experiments in culture were harvested, washed with PBS, centrifuged and delivered as pellets in a sample tube. The pellet was solubilized in denaturation buffer (6M urea/2M thiourea) using 50 µL for every $1x10^6$ cells. Samples are kept chilled on ice, though it should be noted that the concentrated urea solution tends to precipitate upon cooling. Vortex vigorously to break up large pieces, then use an ultrasonic homegenizer and pulse 8-10 times with the maximum output. This improves solubility as well as shearing genomic DNA into smaller fragments to lower sample viscosity. Vortex samples again prior to centrifugation to pellet any cell debris and remaining insoluble material. Transfer solution to a fresh tube and assay protein concentration by Bradford.

7.2.4.1 Wessel-Fluegge extraction of cell lysate.

Wessel-Fluegge extraction is a liquid-liquid extraction technique designed to fractionate biomolecules between immiscible methanol/water and chloroform[122]. Small molecule metabolites and lipids partition to the polar and non-polar phases respectively, while protein precipitates out and collects at the interphase region. This form of extraction is effective at removing small amounts of detergent that may be

present in the sample, as well as removing undesired background molecules that may interfere with the protein analysis. Following the ultrasonic homogenizer step in the previous protocol, add 400 µL each of chilled methanol, water and chloroform to the sample. Place on a shaker at 4 C and shake vigorously for 20 min. Centrifuge samples for 10 min at maximum speed to separate the phases. Remove the polar upper phase with a pipet, avoiding disruption of the interphase region where a white protein precipitate should be visible. Add 600 µl chilled methanol and shake again for 10 min. Centrifuge again to pellet precipitate and decant off supernatant carefully. Air/dry the pellet for 5-10 min in a fume hood, then solubilize in denaturation buffer and assay protein content by Bradford.

7.2.5 Bradford determination of protein concentration.

Dilute protein solution 10 times with water to ensure readings remain within linear range of detector. Combine 2 µL of sample with 498 uL of water and 500 µL of Bradford reagent in a cuvet. Vortex briefly. Measure absorbance at 595 nm after blanking spectrophotometer. Measure absorbance of a serial dilution of bovine serum albumin (BSA) spanning 0.1 to 2 mg/mL and use the linear regression analysis to convert sample absorbances to protein concentrations.

7.2.6 NanoDrop determination of peptide concentration.

A protein sample subjected to a typical proteomic workflow of reduction, alkylation, digestion and desalting is likely to undergo considerable loss of material due to a number of factors, including non-specific binding effects and incomplete C18 capture. It is desirable to measure the concentration of peptides in a sample eluted from a StageTip that is destined to be injected into an LC-MS/MS system to control for either under or or overloading. This can be achieved with the NanoDrop using only 1 µL of sample by measuring the absorbance of the peptide bond at 205 nm with an extinction coefficient of 31 L mol^{-1} cm^{-1}.

7.2.7 SDS-PAGE preparatory gel prior to in-gel enzymatic digestion.

For samples derived from immunoprecipitations where the eluent contains detergents, such as Laemmli sample buffer that has been employed to effect the elution, embedding the proteins in a polyacrylamide matrix provides a means of immobilizing them for the purpose of washing away these MS-incompatible contaminants. Prepare a single-phase 8% gel in a 1.5 mm slab with a 5-position comb. Load sample (up to 80-100 µL volume) and run at a fixed amperage of 55 mA for approximately 10 minutes, time for the sample to migrate completely into the gel and penetrate 1-2 mm. Stain gel for 30 min with Coomassie Plus Imperial stain. Destain with deionized water overnight with multiple changes.

7.2.8 In-gel protein digest.

After the samples have been run on 8% SDS-PAGE preparatory gels, stained and destained, the bands are excised with a razor blade and cut into cubes of approximately 1-2 mm. The gel pieces are washed with 1 mL volumes of 1:1 ethanol/50 mM ABC, 50 mM ABC, and again with 1:1 ethanol/ABC for 20 minutes

each with vigorous shaking, with the aspirated to waste between each step. The pieces are then dehydrated in 1 mL ethanol for 20 min and the solvent aspirated to waste before being placed under vacuum (SpeedVac) for 10 min to remove any remaining ethanol. The dried gel pieces are rehydrated with 250 µL 10 mM dithithreitol (DTT) and incubated at 56 °C for 45 min to reduce disulfide bonds. The solution is removed and replaced with 250 µL 55 mM 2-chloroacetamide (CAA), the gel pieces protected from light and incubated for 45 min at room temperature. The CAA solution is aspirated to waste and the gel pieces washed with 1 mL volumes of 50 mM ABC and 1:1 ethanol:ABC. The gel pieces are dehydrated again with 1 mL ethanol and dried in the SpeedVac before being rehydrated in 300 µL 50mM ABC with 1.5 µg of sequencing-grade trypsin and incubated overnight (16 h) at 30 °C. After digestion, the supernatant is collected in a fresh tube and the gel pieces subjected to extraction with 250 µL volumes of 30% ACN/3% TFA and ACN to retrieve the peptides. The combined volumes of 750 µL are placed in the SpeedVac and the volume reduced to 100-150 µL before being collected and desalted on a C18 StageTip device.

7.2.9 In-solution protein digest.

Protein extract solutions derived from Wessel-Fluegge extraction of MEF cell pellets and assayed by Bradford for protein concentration are subjected to an in-solution digest of proteins. Once the required volume of protein solution equivalent to 120 µg of protein is placed in a fresh 0.5 mL Eppendorf tube, denaturation buffer is used to adjust volumes of all samples in the set (e.g. time course) so that they are equivalent, such that all subsequent volumes and sample handling steps are common. SpikeTides are prepared as a 25 fmol/µL working stock solution and 4.8 uL added to each sample. Tris(2-carboxyethyl)phosphine (TCEP) is added to a final concentration of 5 mM and the samples incubated for 30 min at room temperature to reduce disulfide bonds. CAA is added to a final concentration of 15 mM and the samples incubated for 45 min in the dark. Endoproteinase Lys C is added to each sample at a ratio of 1:40 (w/w) and incubated at 30 °C for 2 h. A volume of 50 mM ABC equivalent to three times the current total volume is added to reduce the urea concentration below 2 M, and then sequencing grade trypsin is added at a ratio of 1:40 (w/w) and the samples incubated at 30 °C overnight (16 h). The sample solutions are acidified by adding 10% TFA to a final concentration of 1% and then subjected to a StageTip extraction, with the volume for a single 120 ug sample divided across six three-disc StageTips.

7.2.10 LC-SRM-MS analysis of NF-kB in MEF cells.

Following the in-solution digestion, the peptides were eluted from the StageTip with 50 µL of 80% ACN/0.1% formic acid into a 96-well plate. The eluate was evaporated under vacuum in a SpeedVac and the sample reconstituted with 20 µL 3% ACN/0.1% formic acid. An Eksigent nanoLC Ultra1D pump coupled with a CTC-PAL HTC-xt autosampler were used to inject 5 µL of sample and resolve it on the self-made reversed-phase analytical column emitter with a gradient from 3% to 36% ACN over 38 minutes at a flow-rate of 250 nL/min. The total gradient length including equilibration time was 60 min, chosen to keep the overall measurement of the complete batch of samples as short as possible to prevent a technical failure of the MS or column from interrupting the analysis. The column eluent was electrosprayed into a Sciex 5500 Q-trap triple-quadrupole mass spectrometer operating at an

ionization potential of 2.3 kV and an interface heater temperature of 150 °C, with a sheath gas setting (GS1) of 15 and a curtain gas setting of 40 to minimize source contamination. The declustering potential (DP) was maintained at 150 V for all SRM transitions, with a . The resolution of quadrupoles 1 and 3 was kept at “Unit”, a peak half-width of 0.7 Da. All samples were measured in triplicate, and the raw data files in .wiff format were analyzed with Sciex MultiQuant v1.1. Using this software, the peaks corresponding to the selected peptides were integrated manually, drawing a baseline between the valleys bounding it on each side. The data was unsmoothed to prevent loss of intensity and to make the peak boundaries more apparent. The peak area for each transition was divided by the peak area for the corresponding transition from the reference SpikeTide to give a ratio. The ratios for all transitions and all injections were averaged and a standard deviation calculated. Five fmols of reference peptide were injected with each sample, and this value was used to convert peak area ratios into absolute values. The absolute values were normalized to the amount of initial total protein represented by the injection, 5 µg, to allow comparison between the wild-type MEF cells and the mutant cell lines which had variable cell volumes. The peptide signals representing the N-terminal products, p50 and p52, were adjusted by subtracting the signal from the C-terminal peptide, which represent the intact p105 and p100, respectively.

7.2.11 LC-MS/MS measurement of peptides derived from in-gel digestion of immunoprecipitations, and bioinformatic analysis of raw data.

Following in-solution or in-gel digestion of proteins derived from whole cell lysates or immunoprecipitations, peptides were eluted from the StageTip with 50 µL of 80% ACN/0.1% formic acid into a 96-well plate. The eluate was evaporated under vacuum in a SpeedVac and the sample reconstituted with 20 µL 3% ACN/0.1% formic acid. A Proxeon nLC nanoflow HPLC with integrated autosampler was used to inject 5 µL of sample and form a gradient from 3 to 38% ACN over 118 min at a flow-rate of 250 nL/min, with a total gradient legth of 155 min including column equilibration. The column eluent was electrosprayed directly into the Q-Exactive operating at an ionization potential of 1.8kV. A Top 10 data-dependent acquisition method was employed, collecting MS1 fullscan spectra from 300 to 1800 m/z at a resolution of 70, 000, an automatic gain control (AGC) target of 1e6 ions and a maximum injection time of 120 ms. The ten most abundant peptide precursor ions were selected for MS/MS via an isolation window of 2 m/z and fragmented by Higher-energy collisional dissociation (HCD) at a normalized collision energy of 26 eV. Fragments were collected for 60 ms (120 ms for lower-complexity immunoprecipitations) with an AGC target of 1e5 ions and scanned by the Orbitrap at a resolution of 17,500 between 200 and 2000 m/z. Once measured, precursor ions were excluded from MS/MS selection for 30 s to avoid redundancy.

Data files from the measurements in .raw format were analyzed by MaxQuant v1.2.2.5 [218] for the NEMO IP samples, and by v1.5.2.8 otherwise. Samples subjected to SILAC labeling used a multiplicity of 2 with Lys8 and Arg10 labels selected. Methionine oxidation and protein N-terminal acetylation were chosen as variable modifications and carbamidomethylation of cysteine residues as a fixed modification. The maximum number of missed cleavages was set at 2 and the digestion mode was set to Trypsin/P. Unlabeled samples measured in triplicate were subjected to a label-free quantification by selecting the LFQ settings with a minimum ratio count of 2. Human samples were analyzed against Uniprot database HUMAN.2014-10.fasta and

mouse samples against Uniprot database MOUSE.2014-10.fasta. Fragment mass tolerance was set to 20 ppm and the false-discovery rate set to 1%. In all cases, match between runs with a window of 0.7 min was enabled to increase coverage. Visualization of data in the output file Protein Groups was achieved using the open-source R Project for Statistical Computing.

7.2.12 In-house manufacture of reversed-phase analytical columns with emitters.

The analytical columns used for all LC-MS measurements were prepared from fused-silica capillary with a polyimide coating from Polymicro Industries and had an outer diameter of 365 µm and an inner diameter of 75 µm. 50 cm lengths of tubing were cut and the polyimide coating in the center 2 cm removed by briefly burning with a flame and rubbing away with an isopropanol-wetted laboratory wipe. The emitters were formed using a Laser Puller P-2000 machine (Sutter Instruments) that softens the fused-silica with a laser and applies force along the capillary's axis to stretch the tubing into a fine point. A Heat setting of 275, a Velocity of 6, and a Delay of 255 ms were used to pull the 50 cm length of tubing into two 25 cm empty columns. The columns were packed with Dr. Maisch Reprosil-Pur C18 material that had a particle size of 3 µm and internal pores of 100 Å using a pressure-injection cell (Biostep, GmbH). After packing the column bed length was cut to 20 cm and the column stored in methanol until use.

7.3 Collaborator contributions: cell culture, immunoprecipitations, and mathematical modelling.

Dr. Buket Yilmaz of the Max Delbrueck Centre for Molecular Medicine (MDC) performed cell culture of MEF cells, including stimulation time courses, Dynamic SILAC labeling, inhibition of ribosome, proteasomes and p97, and analyzed samples in parallel by Western blotting. Mathematical modeling, including calculations and simulations were performed by Dr. Bente Kofahl, Dr. Katherina Baum and Dr. Jana Wolf, with relevant details published in Yilmaz *et al*, Cell Reports, 2014 [25]. Inbal Ipenberg (MDC) labeled MEF cells with SILAC amino acids and performed p100 immunoprecipitations according to the protocol established by Dr. Yilmaz. Nadine Mikuda (MDC) performed cell culture of HepG2, MEF and U20S cells and immunoprecipitations for the interactomics investigations of IKKγ/NEMO . Anne-Kathrin Mueller-Rischart of the Ludwig Maximilians University performed all cell culture of MEF cells for the investigations into Parkin, including pull-downs of transfected cell lines as well as RNA silencing of wild-type MEFs [130]. Rebecca Eccles (MDC) performed cell culture, transfections of HEK293T and NGP neuroblastoma cells for the investigations into PKA inhibition by ARHGAP36 [187].

8 Appendices

8.1 Abbreviations

AAA	ATPaSE associated with diverse cellular activities
AIC_c	Akaike information criterion
ARD	Ankyrin repeat domain
AUC	Area under curve
cAMP	Cyclic adenosine monophosphate
CE	Collision energy
CI	Chemical ionization
CID	Collision-induced dissociation
CoIP	Protein complex immunoprecipitation
DBeQ	N2,N4-dibenzylquinazoline-2,4-diamine
DDA	Data-dependent analysis
DNA	Deoxyribonucleic acid
DP	Declustering potential
EDC4	Enhancer of mRNA decapping protein 4
EI	Electron impact
EMS	Enhanced MS
ER	Enhanced resolution
ERAD	Endoplasmic reticulum associated degradation
ESI	Electrospray ionization
FAB	Fast-atom bombardment
Fmoc	Fluorenylmethyloxycarobonyl
H-Arg-10	Heavy arginine $^{13}C_6{}^{15}N_4$
H-Lys-8	Heavy lysine $^{13}C_6{}^{15}N_2$
HPLC	High-performance liquid chromatography
IBAQ	Intensity-based absolute quantitation
IKK	Inhibitor of κB kinase
IKKY	Inhibitor of nuclear factor kappa-B kinase subunit gamma
K-GG	Lysine with diglycyl remnant from ubiquitin
KO	Knockout
L-Arg-0	Light arginine $^{12}C_6{}^{14}N_4$
L-Lys-0	Light lysine $^{12}C_6{}^{14}N_2$
LC-MS/MS	Liquid chromatography tandem mass spectrometry

LFQ	Label-free quantitation
LT	Lymphotoxin
LTβR	Lymphotoxin beta receptor
LUBAC	Linear ubiquitin assembly complex
M-Arg-6	Medium-heavy arginine $^{13}C_6{}^{14}N_4$
MEF	Murine embryonic fibroblast
MS	Mass spectrometry
MS/MS	Tandem mass spectrometry
NaF	Sodium fluoride
NaV	Sodium vanadate
NEM	N-ethylmaleimide
NF-kB	Nuclear factor κB
NGP	Neuroglioblastoma
NIK	NF-κB inducing kinase
PFL	Protein frequency library
PINK1	PTEN-induced putative kinase protein 1
PKA	Protein kinase A
PKAC	Protein kinase A catalytic subunit
PKAR	Protein kinase A regulatory subunit
PKI	Protein kinase inhibitor
Q1,2 or 3	Quadrupole 1, 2 or 3
QQQ	Triple quadrupole
qRT-PCR	Quantitative real-time polymerase chain reaction
RHD	Rel homology domain
RIP1	Receptor interacting protein kinase 1
RNAseq	Ribonucleic acid sequencing
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel eletrophoresis
SILAC	Stable isotope labeling in cell culture
siRNA	Small interfering RNA
SRM	Selected reaction monitoring
TFA	Trifluoroacetic acid
TNF	Tumor necrosis factor
TOF	Time-of-flight
TRAF	TNF receptor associated factor
Ub	Ubiquitin
UBAN	Ubiquitin binding in ABIN and NEMO

UBL	Ubiquitin-like
UFD	Ubiquitin fusion degradation
VCP/p97	Vasolin-containing protein
WT	Wild-type

8.2 List of Figures

1.1 Signaling via the noncanonical NF-κB pathway. **6**
1.2 Western blot analysis of LTβR-induced MEF cells. **7**
2.1 Solid-phase peptide synthesis using a Rink amide support. **20**
2.2 Schematic representation of triple-quadrupole (QQQ) mass spectrometer. **21**
2.3 Enhanced MS spectrum of the p100 peptide TPSPSGSLLR. **22**
2.4 Enhanced Resolution spectrum of the p100 peptide TPSPSGSLLR. **23**
2.5 Schematic of a generic peptide. B. MS/MS spectra of peptide TPSPSGSLLR. **24**
2.6 Declustering potential optimization. **25**
2.7 Collision energy optimization for p100 peptide TPSPSGSLLR. **26**
2.8 Enhanced MS scan of p105 peptide VGADLSLLDR. **27**
2.9 Enhanced resolution scan of p105 peptide VGADLSLLDR. **27**
2.10 Product ion (MS/MS) spectra of p105 peptide VGADLSLLDR. **28**
2.11 Collision energy optimization for p105 peptide VGADLSLLDR. **28**
2.12 Schematic of tryptic peptides **29**
2.13 Signal response curves for NF-κB peptides. **31**
2.14 SRM-MS measurement of H-Arg-10 p100 peptide in wild-type MEF. **35**
2.15 SRM-MS measurement of L-Arg-0 p100 peptide in wild-type MEF. **36**
2.16 SRM-MS measurement of L-Arg-0 p100 peptide in $p100^{-/-}$ MEF. **36**
2.17 SRM-MS measurement H-Arg-10 p52 peptide in $Nfkb2^{-/-}$ MEF. **37**
2.18 SRM-MS measurement of L-Arg-0 p52 peptide in $Nfkb2^{-/-}$ MEF. **38**
2.19 SRM-MS measurement of H-Arg-10 p105 peptide in wild-type MEF. **38**
2.20 SRM-MS measurement of L-Arg-0 p105 peptide in $p105^{-/-}$ MEF. **39**
2.21 SRM-MS measurement of H-Arg-10 p50 peptide in $Nfkb1^{-/-}$ MEF. **39**
2.22 SRM-MS measurement of L-Arg-0 p50 peptide in $Nfkb1^{-/-}$ MEF. **40**
2.23 SRM-MS measurement of L-Arg-0 p50 peptide in wild-type MEF after Wessel-Fluegge extraction. **40**
2.24 Absolute quantitation of p100, p52, p105 and p50 in MEF cells. **41**
2.25 Absolute quantitation of p100, p52, p105 and p50 in LT-stimulated wild-type MEF cells by SRM-MS. **43**
2.26 Basal and LTβR-stimulated concentration values from the absolute quantitation of NF-κB precursors and products in wild-type MEF cells. **44**
2.27 Workflow schematic of the Dynamic SILAC LT stimulation experiment. **45**

2.28 Quantitation of NF-κB precursors and products in the Dynamic SILAC LT-stimulation time course. **47**
2.29 Dynamic SILAC SRM-MS evaluation of the influence of p97 ATPase activity on M-Arg-6 precursor processing. **48**
2.30 Effect of ribosome or proteasome inhibition on NF-κB precursor processing. **49**
2.31 Proposed mathematical models of NF-κB precursor complex processing. **52**
2.32 Calculated Akaike weights for Models M1 to M4. **53**
2.33 Model simulations of LTβR stimulation by the model M3. **54**
2.34 Incorporation of H-Arg-10 and H-Lys-8 residues in MEF cells by SILAC. **63**
2.35 SILAC p100 interactome Christmas tree plot. **65**
2.36 LFQ p100 interactome volcano plot. **66**
2.37 Incorporation of H-Arg-10 and H-Lys-8 residues into the MEF proteome. **68**
2.38 Radiation-inducible interactors of NEMO. **71**
2.39 Gene ontology analysis of the inducible interactors of NEMO. **72**
2.40 Parkin promotes the linear ubiquitination of NEMO. **76**
2.41 Parkin is a component of LUBAC and binds preferentially to HOIP. **77**
2.42 Coexpression of PKAC and ARHGAP36 results in increased levels of K63-linked polyubiquitination. **79**
2.43 Intensity-based quantification shows ARHGAP36 and PKAC found in equimolar amounts in neuroglioblastoma cells. **80**

8.3 List of Tables

1.1 Comparison of triple-quadrupole and Orbitrap mass analyzers. **12**
2.1 SRM transitions for MS analysis of p100/p52 and p105/p50 protein species. **29**
2.2 SRM method parameters for NF-κB pathway proteins. **57**
2.3 Summary of IKKγ/NEMO IP optimization. **69**
2.4 Number of peptides of known IKKγ/NEMO interactors found in MS screens. **70**
2.5 Polyubiquitin linkage specific SRM methods adapted from literature. **73**

8.4 Publications

Mikuda N, Kolesnichenko M, Beaudette P, Uyar B, Sun W, Alkalin A, Chen W, Dittmar G, Hinz M, Scheidereit C. The IκB kinase complex is a regulator of mRNA stability in response to stress. (submitted)

Eccles RL, Czajkowski MT, Barth C, Müller PM, McShane E, Grunwald S, Beaudette P, Mecklenburg N, Volkmer R, Zühlke K, Dittmar G, Selbach M, Hammes A, Daumke O, Klussmann E, Urbé S, Rocks O. Bimodal antagonism of PKA signalling by ARHGAP36. *Nature Communications*. 2016, 7: 12963

Beaudette P, Popp O, Dittmar G. Proteomic techniques to probe the ubiquitin landscape. *Proteomics*. 2015, 16 (2): 273-287.

Jerke U, Hernandez DP, Beaudette P, Korkmaz B, Dittmar G, Kettritz R. Neutrophil serine proteases exert proteolytic activity on endothelial cells. *Kidney International*. 2015, 8 (4): 764-775.

Yilmaz ZB, Kofahl B, Beaudette P, Baum K, Ipenberg I, Weih F, Wolf J, Dittmar G, Scheidereit C. Quantitative dissection and modelling of the NF-κB p100-p105 module reveal interdependent precursor proteolysis. Cell Reports. 2014, (9): 1756-1769.

Müller-Rischart AK, Pilsl A, Beaudette P, Patra M, Hadian K, Funke M, Peis R, Deinlein A, Schweimer C, Kuhn PH, Lichtenthaler SF, Motori E, Hrelia S, Wurst W, Trümbach D, Langer T, Krappmann D, Dittmar G, Tatzelt J, Winklhofer KF. The E3 ligase parkin maintains mitochondrial integrity by increasing linear ubiquitination of NEMO. *Molecular Cell.* 2013, 49 (5): 908-921.

Beaudette P., Rossi NA, Huegsen PF, Yu X, Shenoi RA, Doucet A, Overall CM, Kizhakkedathu JN; Development of soluble ester-linked aldehyde polymers for proteomics. *Analytical Chemistry*. 2011, 83 (17): 6500-10.

Beaudette P., Bateman K.P.; Discovery stage pharmacokinetics using dried blood spots. *J. Chromatogr B Analyt Technol Biomed Life Sci*. 2004, 809 (1): 153-8.

8.5 Acknowledgements

I am grateful to my supervisors Prof. Dr. Claus Scheidereit and Dr. Gunnar Dittmar for giving me this tremendous opportunity and for their guidance and encouragement during this period. I have learned so much from you both that will benefit me for years to come, but chief amongst it all is the infectious enthusiasm you have for scientific discovery and for helping me to see the bigger picture.

I had the pleasure of working closely with a number of scientists in AG Scheidereit. Dr. Buket Yilmaz made the precursors project possible through her hard work and exceptional commitment to quality, but more importantly for the care and compassion that come so natural to her and helped me through some of the darker days. I thank Nadine Mikuda and Dr. Michael Hinz for their patience and determination; I gained a number of valuable insights into navigating the intersection of molecular biology and proteomics working on our project. Many thanks also to Inbal Ipenberg for her perseverance in a challenging project as well as her sharp wit. Special thanks to Daniela Keyner for so much help with so many things it is impossible to keep track.

I was blessed with many fantastic colleagues in the Mass Spec Core Facility over the years: Rebekka Migotti, Dr. Rick Scavetta, Dr. Gunther Kahlert, Tamara Kanashova, Dr. Julia Kikuchi, Dr. Daniel Perez Hernandez, Dr. Oliver Popp, Evelyn Ramberger, Alina Dagane, Gülkis Baytek, Lorena Suárez-Artiles, Corinna Friedrichs and Dr. Marieluise Kirchner. This is a truly wonderful group of people and I am lucky to have been part of such a fantastic working environment, filled with valuable discussions, mutual support, friendship and a great deal of laughter and fun. I would also like to thank Dr. Matthias Selbach and his group, our closest partners in the MSCF, both for their camaraderie and their feedback and input into my project, with special thanks to Christian Sommer for sharing his extensive technical knowledge, regular assistance with repairing the instruments, and a charmingly hostile sense of humour.

I am grateful also to my collaborators from outside groups. Dr. Rebecca Eccles from AG Rocks at the MDC and Dr. Anne Kathrin Mueller-Rischart of AG Winklhofer from the Ruhr-Universitaet Bochum. I feel fortunate to have made a contribution to both of these fantastic projects that helped me to explore new biological systems and new experimental techniques.

Finally, I thank my family here and in Canada for their constant love and support. I would like to dedicate this thesis to my son Ruben, for being the sweetest and kindest human being I have ever had the pleasure of knowing.

8.6 Eidesstattliche Erklärung

Ich versichere hiermit, dass ich die vorliegende Dissertation selbstständig und ohne die Hilfe Dritter verfasst habe. Andere als die angegebenen Quellen und Hilfsmittel wurden nicht verwendet. Die den benutzten Quellen wörtlich oder inhaltlich entnommenen Stellen sind als solche kenntlich gemacht. Diese Diplomarbeit hat keiner anderen Prüfungsbehörde vorgelegen.

Berlin, 13.11.2017 _____

Patrick Beaudette

9 References

1. Hayden, M. S. & Ghosh, S. NF-κB in immunobiology. *Nature Publishing Group* **21,** 223–244 (2011).
2. Baldwin, A. S. Regulation of cell death and autophagy by IKK and NF-κB: critical mechanisms in immune function and cancer. *Immunological Reviews* **246,** 327–345 (2012).
3. Ben-Neriah, Y. & Karin, M. Inflammation meets cancer, with NF-κB as the matchmaker. *Nat Immunol* **12,** 715–723 (2011).
4. Courtois, G. & Gilmore, T. D. Mutations in the NF-kappaB signaling pathway: implications for human disease. *Oncogene* **25,** 6831–6843 (2006).
5. Razani, B., Reichardt, A. D. & Cheng, G. Non-canonical NF-κB signaling activation and regulation: principles and perspectives. *Immunological Reviews* **244,** 44–54 (2011).
6. Vallabhapurapu, S. & Karin, M. Regulation and function of NF-kappaB transcription factors in the immune system. *Annual Review of Immunology* **27,** 693–733 (2009).
7. Hinz, M., Arslan, S. Ç. & Scheidereit, C. It takes two to tango: IκBs, the multifunctional partners of NF-κB. *Immunological Reviews* **246,** 59–76 (2012).
8. Hinz, M. & Scheidereit, C. The IκB kinase complex in NF-κB regulation and beyond. *EMBO reports* **15,** 46–61 (2014).
9. Hayden, M. S. & Ghosh, S. Shared Principles in NF-κB Signaling. *Cell* **132,** 344–362 (2008).
10. Liu, F., Xia, Y., Parker, A. S. & Verma, I. M. IKK biology. *Immunological Reviews* **246,** 239–253 (2012).
11. Scheidereit, C. IκB kinase complexes: gateways to NF-κB activation and transcription. *Oncogene* **25,** 6685–6705 (2006).
12. Mitchell, S., Vargas, J. & Hoffmann, A. Signaling via the NFκB system. *Wiley Interdisciplinary Reviews: Systems Biology and Medicine* **8,** 227–241 (2016).
13. Schröfelbauer, B., Polley, S., Behar, M., Ghosh, G. & Hoffmann, A. NEMO ensures signaling specificity of the pleiotropic IKKβ by directing its kinase activity toward IκBα. *Molecular Cell* **47,** 111–121 (2012).
14. Sun, S. C. The noncanonical NF-κB pathway. *Immunological Reviews* **246,** 125–140 (2012).
15. Senftleben, U. *et al.* Activation by IKKalpha of a second, evolutionary conserved, NF-kappa B signaling pathway. *Science* **293,** 1495–1499 (2001).
16. Xiao, G., Harhaj, E. W. & Sun, S. C. NF-kappaB-inducing kinase regulates the processing of NF-kappaB2 p100. *Molecular Cell* **7,** 401–409 (2001).
17. Claudio, E., Brown, K., Park, S., Wang, H. & Siebenlist, U. BAFF-induced NEMO-independent processing of NF-kappa B2 in maturing B cells. *Nat Immunol* **3,** 958–965 (2002).
18. Vallabhapurapu, S. *et al.* Non-redundant and complementary functions of adaptor proteins TRAF2 and TRAF3 in a ubiquitination cascade that activates NIK-dependent alternative NF-κB signaling. *Nat Immunol* **9,** 1364–1370 (2008).
19. Zarnegar, B. J. *et al.* Noncanonical NF-kappaB activation requires coordinated assembly of a regulatory complex of the adaptors cIAP1, cIAP2, TRAF2 and TRAF3 and the kinase NIK. *Nat Immunol* **9,** 1371–1378 (2008).

20. Lo, J. C. *et al.* Coordination between NF-kappaB family members p50 and p52 is essential for mediating LTbetaR signals in the development and organization of secondary lymphoid tissues. *Blood* **107,** 1048–1055 (2006).
21. Solan, N. J., Miyoshi, H., Carmona, E. M., Bren, G. D. & Paya, C. V. RelB cellular regulation and transcriptional activity are regulated by p100. *Journal of Biological Chemistry* **277,** 1405–1418 (2002).
22. Heissmeyer, V., Krappmann, D., Wulczyn, F. G. & Scheidereit, C. NF-kappaB p105 is a target of IkappaB kinases and controls signal induction of Bcl-3-p50 complexes. *The EMBO Journal* **18,** 4766–4778 (1999).
23. Heissmeyer, V., Krappmann, D., Hatada, E. N. & Scheidereit, C. Shared pathways of IkappaB kinase-induced SCF(betaTrCP)-mediated ubiquitination and degradation for the NF-kappaB precursor p105 and IkappaBalpha. *Molecular and Cellular Biology* **21,** 1024–1035 (2001).
24. Orian, A. *et al.* SCF(beta)(-TrCP) ubiquitin ligase-mediated processing of NF-kappaB p105 requires phosphorylation of its C-terminus by IkappaB kinase. *The EMBO Journal* **19,** 2580–2591 (2000).
25. Yılmaz, Z. B. *et al.* Quantitative dissection and modeling of the NF-κB p100-p105 module reveals interdependent precursor proteolysis. *CellReports* **9,** 1756–1769 (2014).
26. Elsasser, S. & Finley, D. Delivery of ubiquitinated substrates to protein-unfolding machines. *Nat. Cell Biol.* **7,** 742–749 (2005).
27. Zhang, Z. *et al.* Ter94 ATPase complex targets k11-linked ubiquitinated ci to proteasomes for partial degradation. *Dev. Cell* **25,** 636–644 (2013).
28. Stolz, A., Hilt, W., Buchberger, A. & Wolf, D. H. Cdc48: a power machine in protein degradation. *Trends in Biochemical Sciences* **36,** 515–523 (2011).
29. Schuberth, C. & Buchberger, A. UBX domain proteins: major regulators of the AAA ATPase Cdc48/p97. *Cell. Mol. Life Sci.* **65,** 2360–2371 (2008).
30. Ye, Y., Meyer, H. H. & Rapoport, T. A. The AAA ATPase Cdc48/p97 and its partners transport proteins from the ER into the cytosol. *Nature* **414,** 652–656 (2001).
31. Ramadan, K. *et al.* Cdc48/p97 promotes reformation of the nucleus by extracting the kinase Aurora B from chromatin. *Nature* **450,** 1258–1262 (2007).
32. Dai, R. M. Involvement of Valosin-containing Protein, an ATPase Co-purified with Ikappa Balpha and 26 S Proteasome, in Ubiquitin-Proteasome-mediated Degradation of Ikappa Balpha. *Journal of Biological Chemistry* **273,** 3562–3573 (1998).
33. Chou, T.-F. *et al.* Reversible inhibitor of p97, DBeQ, impairs both ubiquitin-dependent and autophagic protein clearance pathways. *Proc. Natl. Acad. Sci. U.S.A.* **108,** 4834–4839 (2011).
34. Holman, S. W., Sims, P. F. G. & Eyers, C. E. The use of selected reaction monitoring in quantitative proteomics. *Bioanalysis* **4,** 1763–1786 (2012).
35. Marx, V. Targeted proteomics. *Nature Methods* **10,** 19–22 (2013).
36. Thomson, J. J. Discovery of the electron. *Philosophical Magazine* (1897).
37. Aston, F. W. *Isotopes; Arnold: London, 1922*. (c) Phil. Msg, 1920).
38. Biemann, K. Laying the groundwork for proteomics: mass spectrometry from 1958 to 1988. *Journal of Proteomics* **107,** 62–70 (2014).
39. Biemann, K., Gapp, G. & Seibl, J. *Application of mass spectrometry to structure problems. I. Amino acid sequence in peptides*. (Journal of the American Chemical …, 1959).
40. Hudson, G. & Biemann, K. Mass spectrometric sequencing of proteins. The structure of subunit I of monellin. *Biochemical and Biophysical Research*

Communications **71,** 212–220 (1976).
41. Gerber, G. E. *et al.* Partial primary structure of bacteriorhodopsin: sequencing methods for membrane proteins. *Proc Natl Acad Sci USA* **76,** 227–231 (1979).
42. Khorana, H. G. *et al.* Amino acid sequence of bacteriorhodopsin. *Proc Natl Acad Sci USA* **76,** 5046–5050 (1979).
43. Munson, M. & Field, F. H. *Chemical ionization mass spectrometry. II. Esters*. (Journal of the American Chemical …, 1966).
44. Putney, S. D. *et al.* Primary structure of a large aminoacyl-tRNA synthetase. *Science* **213,** 1497–1501 (1981).
45. Fenn, J. B., Mann, M., Meng, C. K., Wong, S. F. & Whitehouse, C. M. Electrospray ionization for mass spectrometry of large biomolecules. *Science* **246,** 64–71 (1989).
46. Wilm, M. & Mann, M. Analytical Properties of the Nanoelectrospray Ion Source. *Anal. Chem.* **68,** 1–8 (1996).
47. Catherman, A. D., Skinner, O. S. & Kelleher, N. L. Biochemical and Biophysical Research Communications. *Biochemical and Biophysical Research Communications* **445,** 683–693 (2014).
48. Lenz, C. & Dihazi, H. Introduction to Proteomics Technologies. *Methods Mol. Biol.* **1362,** 3–27 (2016).
49. Emmett, M. R. & Caprioli, R. M. Micro-electrospray mass spectrometry: Ultra-high-sensitivity analysis of peptides and proteins. *Journal of the American Society for Mass Spectrometry* **5,** 605–613 (1994).
50. Sandra, K. *et al.* Highly efficient peptide separations in proteomics Part 1. Unidimensional high performance liquid chromatography. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* **866,** 48–63 (2008).
51. Scoble, H. A., Martin, S. A. & Biemann, K. Peptide sequencing by magnetic deflection tandem mass spectrometry. *Biochem. J.* **245,** 621–622 (1987).
52. Kondrat, R. W., McClusky, G. A. & Cooks, R. G. *Multiple reaction monitoring in mass spectrometry/mass spectrometry for direct analysis of complex mixtures*. (Analytical Chemistry, 1978).
53. Hunt, D. F., Yates, J. R., Shabanowitz, J., Winston, S. & Hauer, C. R. Protein sequencing by tandem mass spectrometry. *Proc Natl Acad Sci USA* **83,** 6233–6237 (1986).
54. Hines, W. M. *et al.* Protein identification and protein characterization by high-performance time-of-flight mass spectrometry. *J. Protein Chem.* **17,** 525–526 (1998).
55. Schwartz, J. C. & Jardine, I. in *High Resolution Separation and Analysis of Biological Macromolecules Part A: Fundamentals* **270,** 552–586 (Elsevier, 1996).
56. Makarov, A. Electrostatic Axially Harmonic Orbital Trapping: A High-Performance Technique of Mass Analysis. *Anal. Chem.* **72,** 1156–1162 (2000).
57. and, M. H. & Makarov, A. A. Interfacing the Orbitrap Mass Analyzer to an Electrospray Ion Source. *Anal. Chem.* **75,** 1699–1705 (2003).
58. Zubarev, R. A. & Makarov, A. Orbitrap Mass Spectrometry. *Anal. Chem.* **85,** 5288–5296 (2013).
59. Blattner, F. R. *et al.* The Complete Genome Sequence of Escherichia coli K-12. *Science* **277,** 1453–1462 (1997).
60. Kunst, F. *et al.* The complete genome sequence of the Gram-positive bacterium Bacillus subtilis. *Nature* **390,** 249–256 (1997).
61. Pappin, D. J. C., Hojrup, P. & Bleasby, A. J. Rapid identification of proteins

by peptide-mass fingerprinting. *Current Biology* **3,** 327–332 (1993).
62. Yates, J. R., Speicher, S., Griffin, P. R. & Hunkapiller, T. Peptide Mass Maps: A Highly Informative Approach to Protein Identification. *Anal. Biochem.* **214,** 397–408 (1993).
63. James, P., Quadroni, M., Carafoli, E. & Gonnet, G. Protein identification in DNA databases by peptide mass fingerprinting. *Protein Science* **3,** 1347–1350 (1994).
64. Eng, J. K., McCormack, A. L. & Yates, J. R. An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *Journal of the American Society for Mass Spectrometry* **5,** 976–989 (1994).
65. Link, A. J. *et al.* Direct analysis of protein complexes using mass spectrometry. *Nature Biotechnology* **17,** 676–682 (1999).
66. Aebersold, R. & Mann, M. Mass spectrometry-based proteomics. *Nature* **422,** 198–207 (2003).
67. Smolka, M. B., Zhou, H., Purkayastha, S. & Aebersold, R. Optimization of the Isotope-Coded Affinity Tag-Labeling Procedure for Quantitative Proteome Analysis. *Anal. Biochem.* **297,** 25–31 (2001).
68. Goodlett, D. R. *et al.* Differential stable isotope labeling of peptides for quantitation and de novo sequence derivation. *Rapid Communications in Mass Spectrometry* **15,** 1214–1221 (2001).
69. Regnier, F. E. *et al.* Comparative proteomics based on stable isotope labeling and affinity selection. *J. Mass Spectrom.* **37,** 133–145 (2002).
70. Stewart, I. I., Thomson, T. & Figeys, D. 18O Labeling: a tool for proteomics. *Rapid Communications in Mass Spectrometry* **15,** 2456–2465 (2001).
71. Hsu, J. L., Huang, S. Y. & Chen, S. H. Dimethyl multiplexed labeling combined with microcolumn separation and MS analysis for time course study in proteomics. *ELECTROPHORESIS* **27,** 3652–3660 (2006).
72. Boersema, P. J., Raijmakers, R., Lemeer, S., Mohammed, S. & Heck, A. J. R. Multiplex peptide stable isotope dimethyl labeling for quantitative proteomics. *Nature Protocols* **4,** 484–494 (2009).
73. Thompson, A. *et al.* Tandem Mass Tags: A Novel Quantification Strategy for Comparative Analysis of Complex Protein Mixtures by MS/MS. *Anal. Chem.* **75,** 1895–1904 (2003).
74. Wiese, S., Reidegeld, K. A., Meyer, H. E. & Warscheid, B. Protein labeling by iTRAQ: A new tool for quantitative mass spectrometry in proteome research. *Proteomics* **7,** 340–350 (2007).
75. Ong, S.-E. *et al.* Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Molecular & Cellular Proteomics* **1,** 376–386 (2002).
76. Ong, S.-E. & Mann, M. A practical recipe for stable isotope labeling by amino acids in cell culture (SILAC). *Nature Protocols* **1,** 2650–2660 (2006).
77. Graumann, J. *et al.* Stable isotope labeling by amino acids in cell culture (SILAC) and proteome quantitation of mouse embryonic stem cells to a depth of 5,111 proteins. *Mol. Cell Proteomics* **7,** 672–683 (2008).
78. Geiger, T., Cox, J., Ostasiewicz, P., Wiśniewski, J. R. & Mann, M. Super-SILAC mix for quantitative proteomics of human tumor tissue. *Nature Methods* **7,** 383–385 (2010).
79. Scott, L., Lamb, J., Smith, S. & Wheatley, D. N. Single amino acid (arginine) deprivation: rapid and selective death of cultured transformed and malignant cells. *Br. J. Cancer* **83,** 800–810 (2000).
80. Ono, M. *et al.* Label-free quantitative proteomics using large peptide data

sets generated by nanoflow liquid chromatography and mass spectrometry. *Molecular & Cellular Proteomics* **5,** 1338–1347 (2006).
81. Cox, J. *et al.* Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Mol. Cell Proteomics* **13,** 2513–2526 (2014).
82. Krishnamurthy, D. *et al.* Analysis of the human pituitary proteome by data independent label-free liquid chromatography tandem mass spectrometry. *Proteomics* **11,** 495–500 (2011).
83. Choi, H. *et al.* SAINT: probabilistic scoring of affinity purification–mass spectrometry data. *Nature Methods* **8,** 70–73 (2010).
84. Dunham, W. H., Mullin, M. & Gingras, A. C. Affinity-purification coupled to mass spectrometry: Basic principles and strategies. *Proteomics* **12,** 1576–1590 (2012).
85. Keilhauer, E. C., Hein, M. Y. & Mann, M. Accurate protein complex retrieval by affinity enrichment mass spectrometry (AE-MS) rather than affinity purification mass spectrometry (AP-MS). *Mol. Cell Proteomics* **14,** 120–135 (2015).
86. Alber, F. *et al.* The molecular architecture of the nuclear pore complex. *Nature* **450,** 695–701 (2007).
87. Ward, A. B., Sali, A. & Wilson, I. A. Integrative Structural Biology. *Science* **339,** 913–915 (2013).
88. Ciechanover, A. Intracellular protein degradation: from a vague idea thru the lysosome and the ubiquitin–proteasome system and onto human diseases and drug targeting*. *Cell Death Differ.* **12,** 1178–1190 (2005).
89. Hsu, P. P. *et al.* The mTOR-Regulated Phosphoproteome Reveals a Mechanism of mTORC1-Mediated Inhibition of Growth Factor Signaling. *Science* **332,** 1317–1322 (2011).
90. Sharma, K. *et al.* Ultradeep Human Phosphoproteome Reveals a Distinct Regulatory Nature of Tyr and Ser/Thr-Based Signaling. *Cell Reports* **8,** 1583–1594 (2014).
91. Steger, M. *et al.* Phosphoproteomics reveals that Parkinson's disease kinase LRRK2 regulates a subset of Rab GTPases. *eLife* **5,** e12813 (2016).
92. Peng, J. *et al.* A proteomics approach to understanding protein ubiquitination. *Nature Biotechnology* **21,** 921–926 (2003).
93. Wagner, S. A. *et al.* A Proteome-wide, Quantitative Survey of In Vivo Ubiquitylation Sites Reveals Widespread Regulatory Roles. *Molecular & Cellular Proteomics* **10,** M111.013284–M111.013284 (2011).
94. Rose, C. M. *et al.* Highly Multiplexed Quantitative Mass Spectrometry Analysis of Ubiquitylomes. *Cell Syst* **3,** 395–403.e4 (2016).
95. Skaug, B., Jiang, X. & Chen, Z. J. The role of ubiquitin in NF-kappaB regulatory pathways. *Annu. Rev. Biochem.* **78,** 769–796 (2009).
96. Iwai, K. Diverse ubiquitin signaling in NF-κB activation. *Trends in Cell Biology* **22,** 355–364 (2012).
97. Hershko, A. & Ciechanover, A. The ubiquitin system for protein degradation. *Annu. Rev. Biochem.* **61,** 761–807 (1992).
98. Finley, D., Ciechanover, A. & Varshavsky, A. *Ubiquitin as a central cellular regulator. Cell* **116,** S29–32– 2 p following S32 (2004).
99. Finley, D. Recognition and Processing of Ubiquitin-Protein Conjugates by the Proteasome. *Annu. Rev. Biochem.* **78,** 477–513 (2009).
100. Ordureau, A., Münch, C. & Harper, J. W. Quantifying ubiquitin signaling. *Molecular Cell* **58,** 660–676 (2015).
101. Beaudette, P., Popp, O. & Dittmar, G. Proteomic techniques to probe the

ubiquitin landscape. *Proteomics* **16,** 273–287 (2015).
102. Tokunaga, F. & Iwai, K. [Involvement of LUBAC-mediated linear polyubiquitination of NEMO in NF-kappaB activation]. *Tanpakushitsu Kakusan Koso* **54,** 635–642 (2009).
103. Tokunaga, F. *et al.* Involvement of linear polyubiquitylation of NEMO in NF-kappaB activation. *Nature Publishing Group* **11,** 123–132 (2009).
104. Ghaemmaghami, S. *et al.* Global analysis of protein expression in yeast. *Nature* **425,** 737–741 (2003).
105. Wang, Q., Song, C. & Li, C.-C. H. Molecular perspectives on p97-VCP: progress in understanding its structure and diverse biological functions. *J. Struct. Biol.* **146,** 44–57 (2004).
106. Domon, B. & Aebersold, R. Options and considerations when selecting a quantitative proteomics strategy. *Nature Biotechnology* **28,** 710–721 (2010).
107. Zubarev, R. A. The challenge of the proteome dynamic range and its implications for in-depth proteomics. *Proteomics* **13,** 723–726 (2013).
108. Hortin, G. L. & Sviridov, D. The dynamic range problem in the analysis of the plasma proteome. *Journal of Proteomics* **73,** 629–636 (2010).
109. Michalski, A., Cox, J. & Mann, M. More than 100,000 detectable peptide species elute in single shotgun proteomics runs but the majority is inaccessible to data-dependent LC-MS/MS. *J. Proteome Res.* **10,** 1785–1793 (2011).
110. Nilsson, T. *et al.* Mass spectrometry in high-throughput proteomics: ready for the big time. *Nature Methods* **7,** 681–685 (2010).
111. Aebersold, R. & Mann, M. Mass-spectrometric exploration of proteome structure and function. *Nature* **537,** 347–355 (2016).
112. Tabb, D. L. *et al.* Repeatability and reproducibility in proteomic identifications by liquid chromatography-tandem mass spectrometry. *J. Proteome Res.* **9,** 761–776 (2010).
113. Kusebauch, U. *et al.* Human SRMAtlas: A Resource of Targeted Assays to Quantify the Complete Human Proteome. *Cell* **166,** 766–778 (2016).
114. Yost, R. A. & Enke, C. G. Triple quadrupole mass spectrometry for direct mixture analysis and structure elucidation. *Anal. Chem.* **51,** 1251–1264 (1979).
115. Wells, J. M. & McLuckey, S. A. Collision-induced dissociation (CID) of peptides and proteins. *Meth. Enzymol.* **402,** 148–185 (2005).
116. Picotti, P. & Aebersold, R. Selected reaction monitoring–based proteomics: workflows, potential, pitfalls and future directions. *Nature Methods* **9,** 555–566 (2012).
117. Afzal, V., Huang, J. T.-J., Atrih, A. & Crowther, D. J. PChopper: high throughput peptide prediction for MRM/SRM transition design. *BMC Bioinformatics* **12,** 338 (2011).
118. Searle, B. C., Egertson, J. D., Bollinger, J. G., Stergachis, A. B. & MacCoss, M. J. Using Data Independent Acquisition (DIA) to Model High-responding Peptides for Targeted Proteomics Experiments. *Mol. Cell Proteomics* **14,** 2331–2340 (2015).
119. Story, S. C. & Aldrich, J. V. Preparation of protected peptide amides using the Fmoc chemical protocol. Comparison of resins for solid phase synthesis. *Int. J. Pept. Protein Res.* **39,** 87–92 (1992).
120. MacLean, B. *et al.* Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. *Bioinformatics* **26,** 966–968 (2010).
121. Havlis, J. & Shevchenko, A. Absolute quantification of proteins in solutions

and in polyacrylamide gels by mass spectrometry. *Anal. Chem.* **76,** 3029–3036 (2004).
122. Wessel, D. & Flügge, U. I. A method for the quantitative recovery of protein in dilute solution in the presence of detergents and lipids. *Anal. Biochem.* **138,** 141–143 (1984).
123. Daller, B. *et al.* Lymphotoxin-β receptor activation by lymphotoxin-α(1)β(2) and LIGHT promotes tumor growth in an NFκB-dependent manner. *Int. J. Cancer* **128,** 1363–1370 (2011).
124. Milo, R. What is the total number of protein molecules per cell volume? A call to rethink some published values. *BioEssays* **35,** 1050–1055 (2013).
125. Wiśniewski, J. R., Hein, M. Y., Cox, J. & Mann, M. A 'Proteomic Ruler' for Protein Copy Number and Concentration Estimation without Spike-in Standards. *Molecular & Cellular Proteomics* **13,** 3497–3506 (2014).
126. Hinz, M. *et al.* A cytoplasmic ATM-TRAF6-cIAP1 module links nuclear DNA damage signaling to ubiquitin-mediated NF-κB activation. *Molecular Cell* **40,** 63–74 (2010).
127. Shifera, A. S. Protein-protein interactions involving IKKgamma (NEMO) that promote the activation of NF-kappaB. *J. Cell. Physiol.* **223,** 558–561 (2010).
128. Mirzaei, H. *et al.* Characterizing the connectivity of poly-ubiquitin chains by selected reaction monitoring mass spectrometry. *Mol. BioSyst.* **6,** 2004–11 (2010).
129. Narendra, D., Tanaka, A., Suen, D.-F. & Youle, R. J. Parkin is recruited selectively to impaired mitochondria and promotes their autophagy. *J Cell Biol* **183,** 795–803 (2008).
130. Müller-Rischart, A. K. *et al.* The E3 Ligase Parkin Maintains Mitochondrial Integrity by Increasing Linear Ubiquitination of NEMO. *Molecular Cell* **49,** 908–921 (2013).
131. Henn, I. H. *et al.* Parkin mediates neuroprotection through activation of IkappaB kinase/nuclear factor-kappaB signaling. *J. Neurosci.* **27,** 1868–1878 (2007).
132. Gerlach, B. *et al.* Linear ubiquitination prevents inflammation and regulates immune signalling. *Nature* **471,** 591–596 (2011).
133. Tokunaga, F. *et al.* SHARPIN is a component of the NF-κB-activating linear ubiquitin chain assembly complex. *Nature* **471,** 633–636 (2011).
134. Ikeda, F. *et al.* SHARPIN forms a linear ubiquitin ligase complex regulating NF-κB activity and apoptosis. *Nature* **471,** 637–641 (2011).
135. Hadian, K. *et al.* NF-κB essential modulator (NEMO) interaction with linear and lys-63 ubiquitin chains contributes to NF-κB activation. *J. Biol. Chem.* **286,** 26107–26117 (2011).
136. Taylor, S. S., Ilouz, R., Zhang, P. & Kornev, A. P. Assembly of allosteric macromolecular switches: lessons from PKA. *Nat Rev Mol Cell Biol* **13,** 646–658 (2012).
137. Taylor, S. S. *et al.* Dynamics of signaling by PKA. *Biochim. Biophys. Acta* **1754,** 25–37 (2005).
138. Hegde, A. N., Goldberg, A. L. & Schwartz, J. H. Regulatory subunits of cAMP-dependent protein kinases are degraded after conjugation to ubiquitin: a molecular mechanism underlying long-term synaptic plasticity. *Proc Natl Acad Sci USA* **90,** 7436–7440 (1993).
139. Schwanhäusser, B. *et al.* Global quantification of mammalian gene expression control. *Nature* **473,** 337–342 (2011).
140. Wilkins, M. Proteomics data mining. *Expert Rev Proteomics* **6,** 599–603 (2009).

141. Tyers, M. & Mann, M. From genomics to proteomics. *Nature* **422,** 193–197 (2003).
142. de Godoy, L. M. F. *et al.* Comprehensive mass-spectrometry-based proteome quantification of haploid versus diploid yeast. *Nature* **455,** 1251–1254 (2008).
143. Beck, M. *et al.* The quantitative proteome of a human cell line. *Molecular Systems Biology* **7,** 549–549 (2011).
144. Nagaraj, N. *et al.* Deep proteome and transcriptome mapping of a human cancer cell line. *Molecular Systems Biology* **7,** 548–548 (2011).
145. Richards, A. L., Merrill, A. E. & Coon, J. J. Proteome sequencing goes deep. *Curr Opin Chem Biol* **24,** 11–17 (2015).
146. International Human Genome Sequencing Consortium. Finishing the euchromatic sequence of the human genome. *Nature* **431,** 931–945 (2004).
147. Taylor, C. F. *et al.* The minimum information about a proteomics experiment (MIAPE). *Nature Biotechnology* **25,** 887–893 (2007).
148. Lundberg, E. *et al.* Defining the transcriptome and proteome in three functionally different human cell lines. *Molecular Systems Biology* **6,** 450 (2010).
149. Uhlen, M. *et al.* Proteomics. Tissue-based map of the human proteome. *Science* **347,** 1260419–1260419 (2015).
150. Biggin, M. D. Animal transcription networks as highly connected, quantitative continua. *Dev. Cell* **21,** 611–626 (2011).
151. Tacheny, A., Dieu, M., Arnould, T. & Renard, P. Mass spectrometry-based identification of proteins interacting with nucleic acids. *Journal of Proteomics* **94,** 89–109 (2013).
152. Li, J. J., Bickel, P. J. & Biggin, M. D. System wide analyses have underestimated protein abundances and the importance of transcription in mammals. *PeerJ* **2,** e270 (2014).
153. Beaudette, P. & Bateman, K. P. Discovery stage pharmacokinetics using dried blood spots. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* **809,** 153–158 (2004).
154. Anderson, L. & Hunter, C. L. Quantitative mass spectrometric multiple reaction monitoring assays for major plasma proteins. *Molecular & Cellular Proteomics* **5,** 573–588 (2006).
155. Lange, V. *et al.* Targeted quantitative analysis of Streptococcus pyogenes virulence factors by multiple reaction monitoring. *Mol. Cell Proteomics* **7,** 1489–1500 (2008).
156. Picotti, P., Bodenmiller, B., Mueller, L. N., Domon, B. & Aebersold, R. Full dynamic range proteome analysis of S. cerevisiae by targeted proteomics. *Cell* **138,** 795–806 (2009).
157. Xiang, Y. *et al.* Monitoring a Nuclear Factor- B Signature of Drug Resistance in Multiple Myeloma. *Molecular & Cellular Proteomics* **10,** M110.005520–M110.005520 (2011).
158. Zhao, Y. *et al.* Quantification of Activated NF- B/RelA Complexes Using ssDNA Aptamer Affinity - Stable Isotope Dilution--Selected Reaction Monitoring--Mass Spectrometry. *Molecular & Cellular Proteomics* **10,** M111.008771–M111.008771 (2011).
159. Zhao, Y., Tian, B., Edeh, C. B. & Brasier, A. R. Quantitation of the Dynamic Profiles of the Innate Immune Response Using Multiplex Selected Reaction Monitoring-Mass Spectrometry. *Molecular & Cellular Proteomics* **12,** 1513–1529 (2013).
160. Simicevic, J. *et al.* Absolute quantification of transcription factors during

cellular differentiation using multiplexed targeted proteomics. *Nature Methods* **10,** 570–576 (2013).
161. Towbin, H., Staehelin, T. & Gordon, J. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proc Natl Acad Sci USA* **76,** 4350–4354 (1979).
162. Picotti, P., Aebersold, R. & Domon, B. The implications of proteolytic background for shotgun proteomics. *Molecular & Cellular Proteomics* **6,** 1589–1598 (2007).
163. Cham, J. A., Bianco, L., Barton, C. & Bessant, C. MRMaid-DB: a repository of published SRM transitions. *J. Proteome Res.* **9,** 620–625 (2010).
164. Brusniak, M.-Y. K. *et al.* ATAQS: A computational software tool for high throughput transition optimization and validation for selected reaction monitoring mass spectrometry. *BMC Bioinformatics* **12,** 78 (2011).
165. Mohammed, Y. *et al.* PeptidePicker: a scientific workflow with web interface for selecting appropriate peptides for targeted proteomics experiments. *Journal of Proteomics* **106,** 151–161 (2014).
166. Desiere, F. *et al.* The PeptideAtlas project. *Nucleic Acids Res.* **34,** D655–8 (2006).
167. Farrah, T., Deutsch, E. W. & Aebersold, R. Using the Human Plasma PeptideAtlas to study human plasma proteins. *Methods Mol. Biol.* **728,** 349–374 (2011).
168. de Graaf, E. L., Altelaar, A. F. M., van Breukelen, B., Mohammed, S. & Heck, A. J. R. Improving SRM assay development: a global comparison between triple quadrupole, ion trap, and higher energy CID peptide fragmentation spectra. *J. Proteome Res.* **10,** 4334–4341 (2011).
169. MacLean, B. *et al.* Effect of collision energy optimization on the measurement of peptides by selected reaction monitoring (SRM) mass spectrometry. *Anal. Chem.* **82,** 10116–10124 (2010).
170. Aebersold, R., Burlingame, A. L. & Bradshaw, R. A. Western Blots versus Selected Reaction Monitoring Assays: Time to Turn the Tables? *Molecular & Cellular Proteomics* **12,** 2381–2382 (2013).
171. Ghosh, R., Gilda, J. E. & Gomes, A. V. The necessity of and strategies for improving confidence in the accuracy of western blots. *Expert Rev Proteomics* **11,** 549–560 (2014).
172. Whiteaker, J. R. *et al.* Sequential multiplexed analyte quantification using peptide immunoaffinity enrichment coupled to mass spectrometry. *Mol. Cell Proteomics* **11,** M111.015347–M111.015347 (2012).
173. Gilda, J. E. & Gomes, A. V. Stain-Free total protein staining is a superior loading control to β-actin for Western blots. *Anal. Biochem.* **440,** 186–188 (2013).
174. Eaton, S. L. *et al.* Total protein analysis as a reliable loading control for quantitative fluorescent Western blotting. *PLoS ONE* **8,** e72457 (2013).
175. Murphy, R. M., Larkins, N. T., Mollica, J. P., Beard, N. A. & Lamb, G. D. Calsequestrin content and SERCA determine normal and maximal Ca2+ storage levels in sarcoplasmic reticulum of fast- and slow-twitch fibres of rat. *J. Physiol. (Lond.)* **587,** 443–460 (2009).
176. Bertoni, T. A., Perenha-Viana, M. C. Z., Patussi, E. V., Cardoso, R. F. & Svidzinski, T. I. E. Western blotting is an efficient tool for differential diagnosis of paracoccidioidomycosis and pulmonary tuberculosis. *Clin. Vaccine Immunol.* **19,** 1887–1888 (2012).
177. Gassmann, M., Grenacher, B., Rohde, B. & Vogel, J. Quantifying Western blots: pitfalls of densitometry. *ELECTROPHORESIS* **30,** 1845–1855 (2009).

178. Koller, A. & Wätzig, H. Precision and variance components in quantitative gel electrophoresis. *ELECTROPHORESIS* **26,** 2470–2475 (2005).
179. Mordmüller, B., Krappmann, D., Esen, M., Wegener, E. & Scheidereit, C. Lymphotoxin and lipopolysaccharide induce NF-κB-p52 generation by a co-translational mechanism. *EMBO reports* **4,** 82–87 (2003).
180. Cheong, R., Hoffmann, A. & Levchenko, A. Understanding NF-kappaB signaling via mathematical modeling. *Molecular Systems Biology* **4,** 192 (2008).
181. Basak, S., Behar, M. & Hoffmann, A. Lessons from mathematically modeling the NF-κB pathway. *Immunological Reviews* **246,** 221–238 (2012).
182. Wilm, M. Principles of electrospray ionization. *Mol. Cell Proteomics* mcp.R111.009407 (2011). doi:10.1074/mcp.R111.009407
183. Bhagwat, A. S. & Vakoc, C. R. Targeting Transcription Factors in Cancer. *Trends Cancer* **1,** 53–65 (2015).
184. Rahal, R. *et al.* Pharmacological and genomic profiling identifies NF-κB–targeted treatment strategies for mantle cell lymphoma. *Nature Medicine* **20,** 87–92 (2013).
185. Fuchs, O. Transcription factor NF-κB inhibitors as single therapeutic agents or in combination with classical chemotherapeutic agents for the treatment of hematologic malignancies. *Curr Mol Pharmacol* **3,** 98–122 (2010).
186. Wolf, D. H. & Stolz, A. The Cdc48 machine in endoplasmic reticulum associated protein degradation. *Biochim. Biophys. Acta* **1823,** 117–124 (2012).
187. Eccles, R. L. *et al.* Bimodal antagonism of PKA signalling by ARHGAP36. *Nature Communications* **7,** 12963 (2016).
188. Söderberg, O. *et al.* Direct observation of individual endogenous protein complexes in situ by proximity ligation. *Nature Methods* **3,** 995–1000 (2006).
189. Moreira, J. M. A., Thorsen, S. B., Brünner, N. & Stenvang, J. Proximity probing assays for simultaneous visualization of protein complexes in situ. *Expert Rev Proteomics* **10,** 219–221 (2013).
190. Kahraman, A. *et al.* Cross-link guided molecular modeling with ROSETTA. *PLoS ONE* **8,** e73411 (2013).
191. Kalkhof, S. *et al.* Computational modeling of laminin N-terminal domains using sparse distance constraints from disulfide bonds and chemical cross-linking. *Proteins* **78,** 3409–3427 (2010).
192. Tran, B. Q., Goodlett, D. R. & Goo, Y. A. Advances in protein complex analysis by chemical cross-linking coupled with mass spectrometry (CXMS) and bioinformatics. *Biochim. Biophys. Acta* **1864,** 123–129 (2016).
193. Subbotin, R. I. & Chait, B. T. A Pipeline for Determining Protein–Protein Interactions and Proximities in the Cellular Milieu. *Molecular & Cellular Proteomics* **13,** 2824–2835 (2014).
194. Sinz, A. Divide and conquer: cleavable cross-linkers to study protein conformation and protein–protein interactions. *Analytical and Bioanalytical Chemistry* 1–12 (2016). doi:10.1007/s00216-016-9941-x
195. Giese, S. H., Fischer, L. & Rappsilber, J. A Study into the Collision-induced Dissociation (CID) Behavior of Cross-Linked Peptides. *Mol. Cell Proteomics* **15,** 1094–1104 (2016).
196. Hein, M. Y. *et al.* A human interactome in three quantitative dimensions organized by stoichiometries and abundances. *Cell* **163,** 712–723 (2015).
197. Chen, Z. *et al.* Signal-induced site-specific phosphorylation targets I kappa B alpha to the ubiquitin-proteasome pathway. *Genes & Development* **9,** 1586–1597 (1995).

198. Alkalay, I. *et al.* Stimulation-dependent I kappa B alpha phosphorylation marks the NF-kappa B inhibitor for degradation via the ubiquitin-proteasome pathway. *Proc Natl Acad Sci USA* **92,** 10599–10603 (1995).
199. Kanarek, N., London, N., Schueler-Furman, O. & Ben-Neriah, Y. Ubiquitination and degradation of the inhibitors of NF-kappaB. *Cold Spring Harb Perspect Biol* **2,** a000166–a000166 (2010).
200. Xu, M., Skaug, B., Zeng, W. & Chen, Z. J. A ubiquitin replacement strategy in human cells reveals distinct mechanisms of IKK activation by TNFalpha and IL-1beta. *Molecular Cell* **36,** 302–314 (2009).
201. Matsumoto, M. L. *et al.* K11-linked polyubiquitination in cell cycle control revealed by a K11 linkage-specific antibody. *Molecular Cell* **39,** 477–484 (2010).
202. Dynek, J. N. *et al.* c-IAP1 and UbcH5 promote K11-linked polyubiquitination of RIP1 in TNF signalling. *The EMBO Journal* **29,** 4198–4209 (2010).
203. Newton, K. *et al.* Ubiquitin Chain Editing Revealed by Polyubiquitin Linkage-Specific Antibodies. *Cell* **134,** 668–678 (2008).
204. Carr, S. A. *et al.* Targeted peptide measurements in biology and medicine: best practices for mass spectrometry-based assay development using a fit-for-purpose approach. in **13,** 907–917 (American Society for Biochemistry and Molecular Biology, 2014).
205. Zolg, D. P. *et al.* Building ProteomeTools based on a complete synthetic human proteome. *Nature Methods* **14,** 259–262 (2017).
206. Hanke, S., Besir, H., Oesterhelt, D. & Mann, M. Absolute SILAC for accurate quantitation of proteins in complex mixtures down to the attomole level. *J. Proteome Res.* **7,** 1118–1130 (2008).
207. Zhang, X. Less is More: Membrane Protein Digestion Beyond Urea–Trypsin Solution for Next-level Proteomics. *Molecular & Cellular Proteomics* **14,** 2441–2453 (2015).
208. Addona, T. A. *et al.* Multi-site assessment of the precision and reproducibility of multiple reaction monitoring-based measurements of proteins in plasma. *Nature Biotechnology* **27,** 633–641 (2009).
209. Frattini, A. *et al.* High variability of genomic instability and gene expression profiling in different HeLa clones. *Nature Publishing Group* 1–9 (2015). doi:10.1038/srep15377
210. Chernobrovkin, A. L. & Zubarev, R. A. Detection of Viral Proteins in Human Cells Lines by Xeno-Proteomics: Elimination of the Last Valid Excuse for Not Testing Every Cellular Proteome Dataset for Viral Proteins. *PLoS ONE* **9,** e91433–6 (2014).
211. Michalski, A. *et al.* Mass Spectrometry-based Proteomics Using Q Exactive, a High-performance Benchtop Quadrupole Orbitrap Mass Spectrometer. *Molecular & Cellular Proteomics* **10,** M111.011015–M111.011015 (2011).
212. Gallien, S. *et al.* Targeted Proteomic Quantification on Quadrupole-Orbitrap Mass Spectrometer. *Molecular & Cellular Proteomics* **11,** 1709–1723 (2012).
213. Carr, S. *et al.* The need for guidelines in publication of peptide and protein identification data: Working Group on Publication Guidelines for Peptide and Protein Identification Data. *Molecular & Cellular Proteomics* **3,** 531–533 (2004).
214. Bradshaw, R. A., Burlingame, A. L., Carr, S. & Aebersold, R. Protein identification: the good, the bad, and the ugly. *Molecular & Cellular Proteomics* **4,** 1221–1222 (2005).
215. Guo, X., Trudgian, D. C., Lemoff, A., Yadavalli, S. & Mirzaei, H. Confetti: a multiprotease map of the HeLa proteome for comprehensive proteomics.

Mol. Cell Proteomics **13,** 1573–1584 (2014).
216. Zhang, X. & Miller, K. W. Dodecyl maltopyranoside enabled purification of active human GABA type A receptors for deep and direct proteomic sequencing. *Mol. Cell Proteomics* **14,** 724–738 (2015).
217. Rappsilber, J., Ishihama, Y. & Mann, M. Stop and go extraction tips for matrix-assisted laser desorption/ionization, nanoelectrospray, and LC/MS sample pretreatment in proteomics. *Anal. Chem.* **75,** 663–670 (2003).
218. Cox, J. & Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nature Publishing Group* **26,** 1367–1372 (2008).